



Mesure de la masse du boson de Higgs se désintégrant en quatre leptons dans l'expérience CMS auprès du LHC.

Stéphanie Baffioni

► To cite this version:

Stéphanie Baffioni. Mesure de la masse du boson de Higgs se désintégrant en quatre leptons dans l'expérience CMS auprès du LHC.. Physique des Hautes Energies - Expérience [hep-ex]. Université Paris-Sud, 2015. tel-01284569

HAL Id: tel-01284569

<https://hal.science/tel-01284569>

Submitted on 9 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MÉMOIRE

présenté devant

l'Université Paris-Sud

pour obtenir

l'habilitation à la direction de recherches

par

Stéphanie Baffioni

**Mesure de la masse du boson de Higgs
se désintégrant en quatre leptons dans
l'expérience CMS auprès du LHC.**

Soutenu le 17 décembre 2015, devant la commission d'examen :

M. Olivier	Drapier	<i>rapporteur</i>
M. Arnaud	Lucotte	<i>rapporteur</i>
M. Emmanuel	Perez	<i>rapporteur</i>
M. Ivica	Puljak	<i>président</i>
M. Laurent	Serin	

MÉMOIRE

présenté devant

l'Université Paris-Sud

pour obtenir

l'habilitation à la direction de recherches

par

Stéphanie Baffioni

**Mesure de la masse du boson de Higgs
se désintégrant en quatre leptons dans
l'expérience CMS auprès du LHC.**

Soutenu le 17 décembre 2015, devant la commission d'examen :

M. Olivier	Drapier	<i>rapporteur</i>
M. Arnaud	Lucotte	<i>rapporteur</i>
M. Emmanuel	Perez	<i>rapporteur</i>
M. Ivica	Puljak	<i>président</i>
M. Laurent	Serin	

À Pédro, Nathanaël, et Siméo.

Table des matières

Introduction	5
1 Produire	7
1.1 Collisionner	8
1.1.1 Le boson de Higgs	8
1.1.2 Le LHC	8
1.1.3 Les données	10
1.2 Détecter	12
1.2.1 Désintégration du boson de Higgs	12
1.2.2 Le détecteur CMS	13
1.2.3 Le détecteur de traces (trajectographe)	15
1.2.4 Le calorimètre électromagnétique (ECAL)	18
1.2.4.1 Détection	18
1.2.4.2 Électronique de lecture	20
1.2.4.3 Résolution en énergie	21
1.2.5 Le calorimètre hadronique (HCAL)	22
2 Déclencher	25
2.1 Système de déclenchement	26
2.2 Déclenchement pour les électrons	27
2.2.1 Niveau 1 (L1)	27
2.2.1.1 Chaîne de déclenchement	27
2.2.1.2 Réjection des signaux anormaux	29
2.2.2 Niveau supérieur (HLT)	30
2.3 Déclenchement pour l'analyse du boson de Higgs en quatre leptons	31
2.3.1 Sélection des événements Higgs en quatre leptons	31
2.3.2 Sélection des événements pour les mesures de performance des électrons	34
2.4 Études d'efficacité du <i>trigger</i> diélectron	35
2.4.1 Efficacités par électron	35
2.4.2 Efficacités du <i>trigger</i> diélectron pour les événements Higgs en quatre électrons	39

3	Reconstruire	43
3.1	Défis	44
3.2	Échantillons de données et de simulation	45
3.3	Reconstruction algorithmique des électrons	47
3.3.1	Agrégation d'énergie dans le ECAL	47
3.3.2	Traces	49
3.3.2.1	<i>Seeding</i>	50
3.3.2.2	Reconstruction des traces	55
3.3.3	Agrégation PF de l'énergie des électrons	59
3.3.4	Association trace-cluster	60
3.3.5	Résolution des ambiguïtés	62
3.3.6	Estimation de la charge	62
3.4	Mesure de l'impulsion des électrons	66
3.4.1	Classification	66
3.4.2	Mesure d'énergie dans le ECAL	67
3.4.2.1	Énergies individuelles dans les cristaux	67
3.4.2.2	Corrections de l'énergie du SC	68
3.4.2.3	Échelle et résolution en énergie du SC	71
3.4.3	Combinaison de l'énergie du SC et de l'impulsion de la trace	72
3.4.4	Échelle et résolution en impulsion finales	73
3.5	Sélection des électrons	78
3.5.1	Identification	78
3.5.2	Isolation	84
3.5.3	Réjection des photons convertis	87
3.5.4	Sélection des électrons pour l'analyse $H \rightarrow ZZ^* \rightarrow 4\ell$	88
3.6	Mesures d'efficacité et de probabilité de fausse identification	89
3.6.1	Efficacité de reconstruction	91
3.6.2	Efficacité de sélection	91
3.6.3	Probabilité de fausse identification	95
4	Mesurer	99
4.1	Considérations générales	100
4.1.1	Stratégie	100
4.1.2	Signature et bruits de fond	100
4.1.3	Échantillons de données et de simulation	101
4.2	Sélection des événements	102
4.2.1	Leptons	103
4.2.2	Radiations dans l'état final	103
4.2.3	Sélections cinématiques	104
4.3	Estimation des bruits de fond	105
4.3.1	Bruit de fond irréductible ZZ	105
4.3.2	Bruit de fond réductible	106
4.3.2.1	Stratégie	106

4.3.2.2	Probabilités de fausse identification	107
4.3.2.3	Méthode OS	107
4.3.2.4	Méthode SS	108
4.3.2.5	Résultat	108
4.4	Discriminant cinématique	108
4.5	Petit historique de la découverte	112
4.6	Mesure de la masse	117
4.6.1	Modélisation de $m_{4\ell}$	117
4.6.2	Incertitudes sur $m_{4\ell}$ par événement	118
4.6.3	Discriminant cinématique	120
4.6.4	Incertitudes systématiques	120
4.6.5	Résultats	120
4.6.5.1	Mesure de m_H avec le canal en quatre leptons dans CMS . .	120
4.6.5.2	Combinaison ultime	122
Conclusion		125
Bibliographie		127
Remerciements		137

Introduction

Voici venu un temps rare et précieux dans la vie d'un physicien, dans l'intervalle duquel, avec une once de chance et de volonté, l'attention est susceptible d'être détournée des sollicitations chroniques d'ordre urgent et important afin de poser une question simple : "Qu'ai-je fait durant ces années?". Instantanément, je mesure les chances qui m'ont été données. En premier lieu, celle de vivre l'aventure fabuleuse d'une découverte majeure. Ensuite, celle de mon échelle de temps avant cet aboutissement : assez longue pour apprécier à sa juste valeur la phase de préparation, assez courte pour ne pas subir la sensation de "perdre son temps" ni de "tourner en rond". Enfin, celle de se pencher sur des sujets cohérents dans un environnement intellectuel stimulant et enrichissant.

En ne résistant pas au petit plaisir d'une simple figure, mon humble cheminement en physique des particules sur collisionneur est illustré dans la Fig. 1, montrant l'énergie dans le centre de masse des collisions en fonction des années. Les deux premières transitions attestent du passage de HERA au Tevatron, puis au LHC. Alors que l'évolution semble stable, la brusque descente signe la réalité de l'accident du LHC en 2008, et la remontée illustre les performances et succès du LHC. Il peut être amusant de constater que l'énergie dans le centre de masse prévue et simulée se retrouve systématiquement supérieure ou égale à celle des données. Cette petite parenthèse mise à part, ce manuscrit se concentre sur l'expérience CMS au LHC.

Durant les années depuis mon arrivée au laboratoire Leprince Ringuet (LLR) dans le groupe CMS en 2004, je me suis penchée sur différents sujets, en particulier les tests en faisceaux pour le calorimètre électromagnétique (ECAL), l'émulation des primitives de déclenchement du ECAL, la reconstruction des électrons et le boson de Higgs en quatre leptons. Tous ne sont pas abordés dans ce manuscrit, dont le sujet est la mesure de la masse du boson de Higgs, au travers du canal de désintégration en quatre leptons, avec un point de vue volontairement accentué vers les électrons. Ce document n'est ni une liste exhaustive de mes implications sur CMS, ni une revue générale de la reconstruction des électrons et de l'analyse du boson de Higgs en quatre leptons, mais une mixture des deux.

Quelques libertés sont prises tout au long de ce manuscrit. En particulier, certains termes usuels en anglais ne sont pas traduits, lorsque la traduction est jugée trop lourde, peu esthétique, ou peu pratique. En particulier les mots comme *barrel* (tonneau), *endcap*

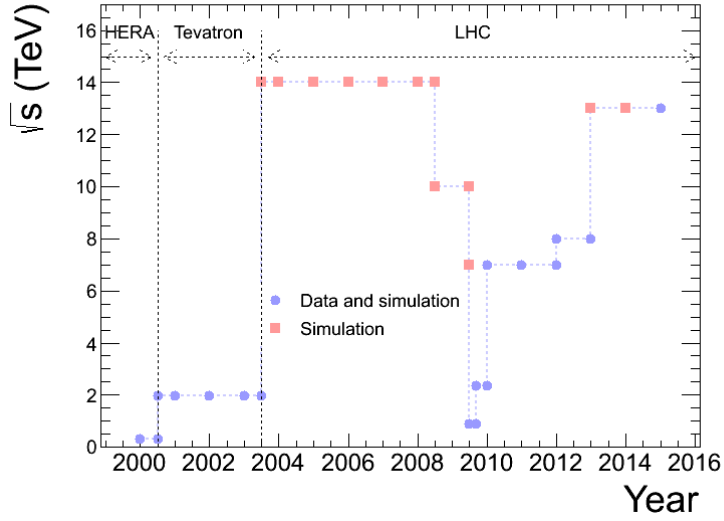


FIGURE 1 – Énergie dans le centre de masse des collisions en fonction des années, durant mon parcours en physique des hautes énergies sur collisionneur.

(bouchon), *cluster* (agrégat), bremsstrahlung (rayonnement continu de freinage), *trigger* (déclenchement), *seed* (graine), *hit* (point de mesure), *run* (unité de prise de données), *blinding* (aveuglement) sont utilisés en anglais et généralement mis en italique. Les nombres sont également écrits dans la forme anglaise, avec des points et non des virgules. En outre, le système d’unités naturelles est utilisé. De plus, le style varie en fonction des chapitres, plutôt léger par endroits, il se fait plus technique à d’autres, à l’instar du chapitre sur la reconstruction des électrons, le plus technique. Enfin, aussi gratifiant qu’eût pu être le contraire, les études réalisées et les résultats exposés sont le fruit de collaborations multiples, voire du travail exclusif d’autres personnes.

La trame de ce document est structurée autour de la notion d’ingrédients, qui, pas à pas, sont nécessaires à la réalisation de la mesure finale. Le premier chapitre, “Produire”, dresse de façon rapide le contexte théorique, puis expérimental, amenant à la production du boson de Higgs. Le deuxième chapitre, “Déclencher”, est consacré à la façon dont les événements sont attentivement choisis pour être enregistrés par CMS. Le chapitre central, “Reconstruire”, décrit de façon détaillée la reconstruction et l’identification des électrons. Enfin le chapitre final, “Mesurer”, est dédié à l’analyse du boson de Higgs en quatre leptons et à la mesure de sa masse.

Il me reste à souhaiter au lecteur un moment agréable et, je l’espère, une lecture intéressante.

Chapitre 1

Produire

Sommaire

1.1	Collisionner	8
1.1.1	Le boson de Higgs	8
1.1.2	Le LHC	8
1.1.3	Les données	10
1.2	Détecter	12
1.2.1	Désintégration du boson de Higgs	12
1.2.2	Le détecteur CMS	13
1.2.3	Le détecteur de traces (trajectographe)	15
1.2.4	Le calorimètre électromagnétique (ECAL)	18
	1.2.4.1 Détection	18
	1.2.4.2 Électronique de lecture	20
	1.2.4.3 Résolution en énergie	21
1.2.5	Le calorimètre hadronique (HCAL)	22

Ce chapitre traite du premier ingrédient nécessaire à la mesure de la masse du boson de Higgs : sa présence. Il aborde sa création au sens large, en effleurant le concept purement théorique pour se concentrer sur sa production concrète au sein du LHC et particulièrement la question de sa détection dans CMS.

1.1 Collisionner

1.1.1 Le boson de Higgs

Idée germée en 1964, le mécanisme de Brout-Englert-Higgs [1–6] est responsable, lorsqu’il est appliqué dans le cadre du modèle standard de la physique des particules et de leurs interactions [7–11], de la brisure spontanée de la symétrie électrofaible impliquant une masse non nulle aux bosons W et Z. Le champ de Higgs, par son interaction avec les fermions, leur permet d’acquérir une masse [12, 13].

Conséquence directe de ce champ de Higgs et pierre angulaire du modèle standard, le boson de Higgs, bien qu’activement recherché depuis, a réussi à échapper pendant 48 ans aux enquêtes les plus minutieuses et les plus approfondies. Il est à noter à la décharge des détectives que pour compliquer l’enquête, son signalement quoique très précis dès que sa masse est connue, ne fait aucun état de celle-ci. Les indications théoriques la contraignent à être ≤ 1 TeV [14–17]. Au fil des années d’investigations, les mesures de précision électrofaibles la limitent à ≤ 152 GeV, à 95 % de niveau de confiance [18]. Durant les deux décennies précédant la fameuse année 2012, les recherches directes au LEP et au Tevatron excluent une masse ≤ 114.4 GeV [19] et entre 162 et 166 GeV [20]. Jusqu’à ce que le LHC intervienne dans l’enquête...

1.1.2 Le LHC

Bien que le champ de Higgs soit supposé omniprésent, il est impossible de croiser un boson de Higgs se baladant dans la rue, ce dernier ayant la bonne idée de se désintégrer sitôt produit dans l’univers, fusse-t-il produit. Curieusement, et comme souvent en physique des particules, le premier ingrédient nécessaire à la recherche de ce boson est donc sa présence, forcée. Cela représente l’un des enjeux majeurs de l’accélérateur et collisionneur LHC (Large Hadron Collider), créé dans le but d’essayer de répondre à certaines des grandes questions de la physique telles que l’origine de la brisure de la symétrie électro-faible, ainsi que les mystères de la matière noire et de l’antimatière.

Parmi les produits de milliards de collisions entre protons, qui, 100 m sous terre dans un vide poussé, parcourent à une vitesse vertigineuse l’anneau accélérateur de 27 km de circonférence, se cachent donc ces fameux bosons responsables de tant d’encre coulée et de cerveaux fusionnés.

Au LHC, les productions principales du boson de Higgs s'effectuent par fusion de gluons, fusion de bosons W ou Z, ou en association avec un W, un Z ou une paire $t\bar{t}$ ou $b\bar{b}$. La Fig. 1.1 représente les sections efficaces de production du boson de Higgs en fonction de sa masse pour une énergie dans le centre de masse égale à 8 TeV.

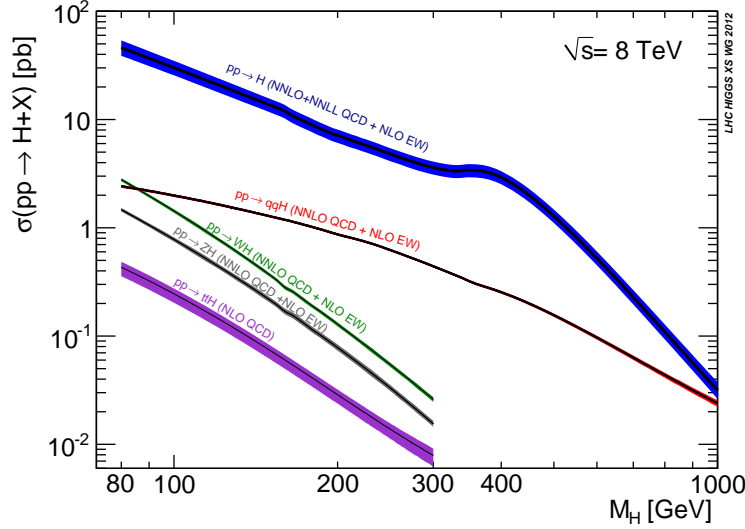


FIGURE 1.1 – Sections efficaces de production du boson de Higgs au LHC en fonction de sa masse, pour une énergie dans le centre de masse égale à 8 TeV.

Cependant, le boson de Higgs est loin d'être la seule particule produite dans les collisions du LHC. Ceci est illustré par la Fig. 1.2 qui montre, en fonction de l'énergie dans le centre de masse, les sections efficaces de production des différents processus du modèle standard. À 8 TeV, le rapport entre la section efficace de production d'un boson de Higgs à 120 GeV [21] et la section efficace totale est de $\approx 5.10^{-10}$, soit ≈ 5 bosons de Higgs produits sur 10 milliards de collisions de protons, d'où la nécessité de produire une quantité de collisions considérable d'une part, et d'effectuer un tri drastique et efficace afin d'enregistrer les données utiles d'autre part. Au jeu des analogies et de l'expression bien connue "rechercher une aiguille dans une botte de foin", cela correspondrait pour un seul boson de Higgs à chercher une aiguille (qui pour les besoins de l'exercice est ici modélisée par un parallélépipède carré de $1 \text{ mm} \times 1 \text{ mm}$ de section droite et de 4 cm de long) dans une botte de foin d'environ 80 m^3 .

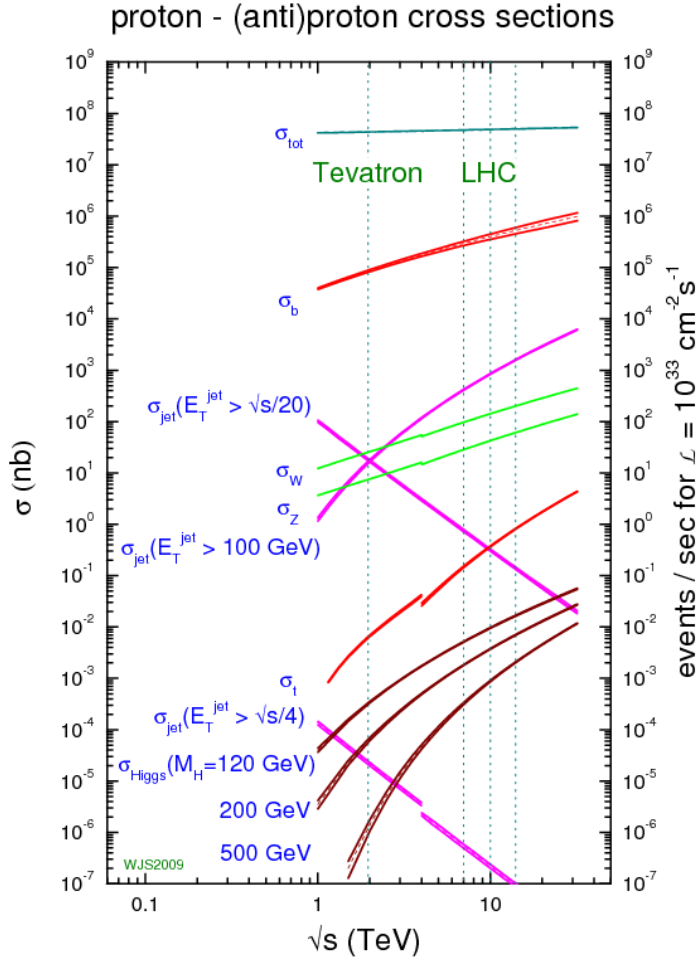


FIGURE 1.2 – Sections efficaces de production en collisions protons-(anti)protons.

1.1.3 Les données

Après un démarrage éclair en 2008, le LHC a opéré ses premières collisions protons-protons fin 2009 à une énergie de 900 GeV, puis de 2.36 TeV dans le centre de masse et à très faible luminosité. Dès lors, le LHC a augmenté et amélioré ses performances de façon considérable pour atteindre 8 TeV dans le centre de masse et une luminosité instantanée égale à $7.7 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ fin 2012, achevant ainsi sa première période de prise de données (Run 1) en dépassant les prévisions les plus optimistes. Les données enregistrées par CMS ont été exploitées dans le cadre de ce mémoire et leurs caractéristiques depuis 2010 sont résumées dans la Table 1.1 et illustrées dans la Fig. 1.3.

Année	\sqrt{s} (TeV)	\mathcal{L}_{tot} (fb^{-1})	$\mathcal{L}_{\text{inst}}^{\text{max}}$ ($\times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$)	$\langle n_{\text{PU}} \rangle$
2010	7	0.036	0.2	2
2011	7	5.1	4.0	9
2012	8	19.7	7.7	21

TABLE 1.1 – Caractéristiques des données enregistrées et certifiées de bonne qualité par CMS. Sont indiqués pour chaque année de prise de données : l’énergie dans le centre de masse (\sqrt{s}), les luminosités intégrée (\mathcal{L}_{tot}) et instantanée maximale ($\mathcal{L}_{\text{inst}}^{\text{max}}$), ainsi que le nombre d’événements supplémentaires moyen par croisement de faisceaux ($\langle n_{\text{PU}} \rangle$). Le temps entre chaque croisement de faisceaux est d’au minimum 50 ns.

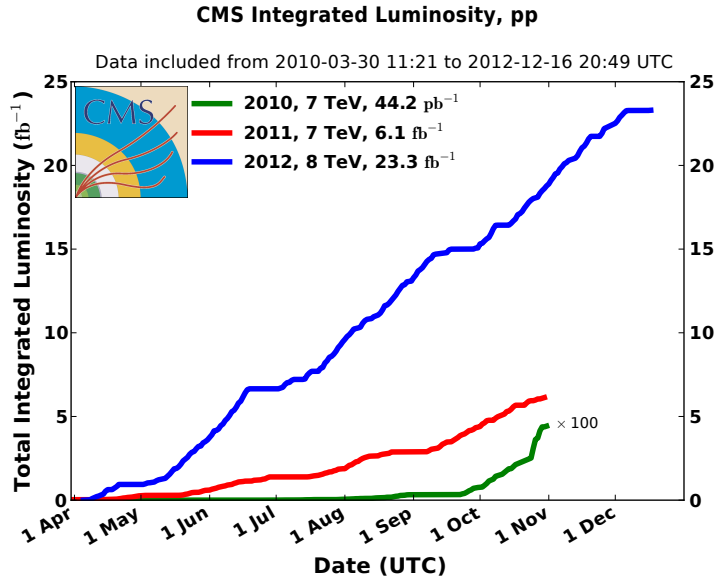


FIGURE 1.3 – Luminosité intégrée délivrée par le LHC à CMS en 2010, 2011 et 2012.

Le premier ingrédient à la recherche du boson de Higgs, sa présence et par conséquent sa production, est donc obtenu grâce aux collisions du LHC. Il faut par la suite être capable de l’attraper, ou de le voir, ou d’en voir la trace, bref de le détecter. Le seul moyen, puisqu’il se désintègre dès sa production, consiste en la détection de ses produits de désintégration. Toute la complexité réside dans le fait que, non content d’être très rare et de masse inconnue, il se désintègre dans de nombreuses voies possibles, dont la proportion change en fonction de sa masse, les canaux les plus identifiables expérimentalement se révélant souvent les plus rares.

1.2 Détecter

1.2.1 Désintégration du boson de Higgs

La Fig. 1.4 illustre cette complexité, et par là-même celle de la recherche du boson de Higgs, en montrant les probabilités de ses différentes désintégrations possibles en fonction de sa masse. Chaque voie de désintégration représente un ou plusieurs canaux d’analyse des données du LHC, certaines voies étant intermédiaires et impliquant l’étude des produits finaux de désintégration, à l’image de celle en deux bosons Z .

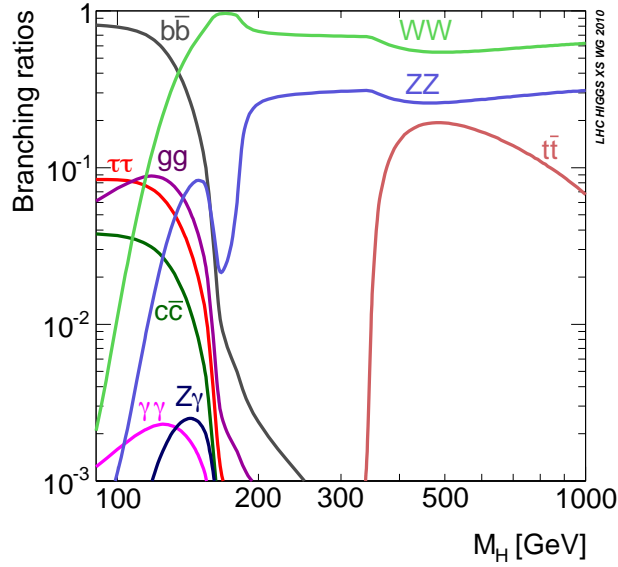


FIGURE 1.4 – Rapports de branchement du boson de Higgs en fonction de sa masse.

Ce mémoire traite du canal leptonique de désintégration du ZZ , $H \rightarrow Z^{(*)}Z^{(*)} \rightarrow 2\ell 2\ell'$, $\ell = e, \mu$. Il présente plusieurs avantages expérimentaux majeurs. Tout d’abord, ce canal possède une signature très “propre” dans le détecteur, c’est-à-dire peu de particules produites, laissant chacune peu de traces ou dépôts dans les détecteurs, un avantage de taille dans les événements issus de collisions hadroniques. Par ailleurs, toutes les particules issues de la désintégration du boson de Higgs peuvent être détectées, ce qui permet d’accéder à des informations fondamentales telles que sa masse et ce de façon très précise, la mesure des leptons e et μ offrant la meilleure résolution possible par rapport à des objets plus complexes tels que les jets ou les leptons τ . Enfin, ce canal bénéficie d’un niveau de bruit de fond remarquablement bas. Tous ces avantages en font l’un des canaux les plus sensibles pour la recherche du boson de Higgs au LHC, voire le canal le plus sensible dans une large gamme de masse. La contrepartie de ces avantages réside dans la rareté du signal, comme illustré dans la Fig. 1.5 montrant les probabilités des différentes désintégrations finales comprenant des leptons en fonction de la masse du boson de Higgs (jusqu’à 250 GeV). Dans le cas d’une

masse du boson de Higgs comprise entre 120 et 130 GeV, le rapport d'embranchement en 4 leptons est de l'ordre de 0.01 %. Cette rareté implique de préserver l'efficacité la plus élevée possible dans toutes les étapes de l'analyse, en particulier dans la détection, la reconstruction et l'identification des leptons, dont la moindre défaillance en efficacité se propage en puissance 4 sur les événements sélectionnés.

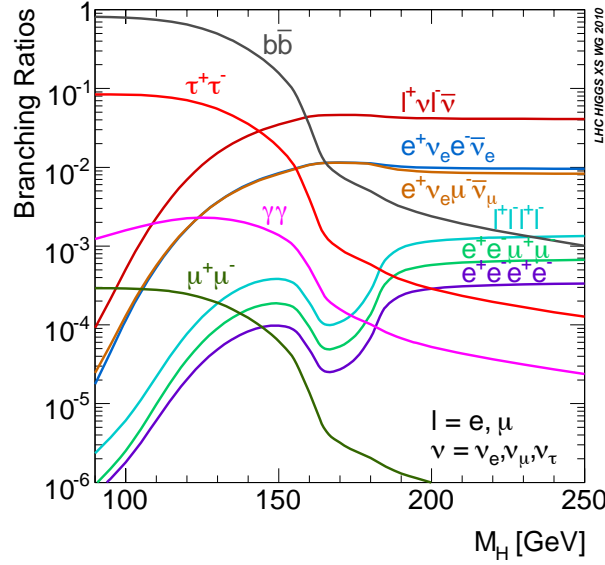


FIGURE 1.5 – Rapports de branchement du boson de Higgs en leptons en fonction de sa masse. Le rapport de branchement en $b\bar{b}$ est montré à titre comparatif.

Une fois les bosons de Higgs produits et sachant par quels produits de désintégration le chercher, le prochain ingrédient est la détection de ces produits. Les paragraphes suivants décrivent le détecteur CMS, et en particulier les sous-parties permettant la détection des électrons (employé de façon générique pour désigner des électrons ou des positrons), sur lesquels ce mémoire se concentre principalement.

1.2.2 Le détecteur CMS

Le détecteur CMS (Compact Muon Solenoid), installé autour d'un des points de collisions des faisceaux du LHC, est un ensemble cylindrique de 14500 tonnes, 21.6 m de longueur et 14.6 m de diamètre [22]. Il est composé de sous-détecteurs agencés de part et d'autre de l'élément central, le solénoïde supraconducteur. Celui-ci, d'un diamètre interne égal à 6 m et d'une longueur de 13 m, délivre un champ magnétique de 3.8 Tesla impliquant la courbure et ainsi la mesure des impulsions des particules chargées produites lors des collisions, et ce jusqu'à des impulsions de l'ordre du TeV, tout en permettant un détecteur compact.

À l'intérieur, sont placés le trajectographe, le calorimètre électromagnétique (ECAL) et les détecteurs de pied de gerbe pour les parties à l'avant (*preshowers*), et le calorimètre hadronique (HCAL). À l'extérieur, insérés dans la culasse de l'aimant, des détecteurs gazeux à ionisation permettent la mesure des muons. Des calorimètres placés à l'avant permettent l'extension de la couverture de détection. La Fig. 1.6 illustre la structure de CMS. Les électrons laissant une trace dans le trajectographe et déposant ensuite leur énergie dans le ECAL, ces deux détecteurs sont décrits en détails dans les prochains paragraphes, avec également quelques informations sur le HCAL.

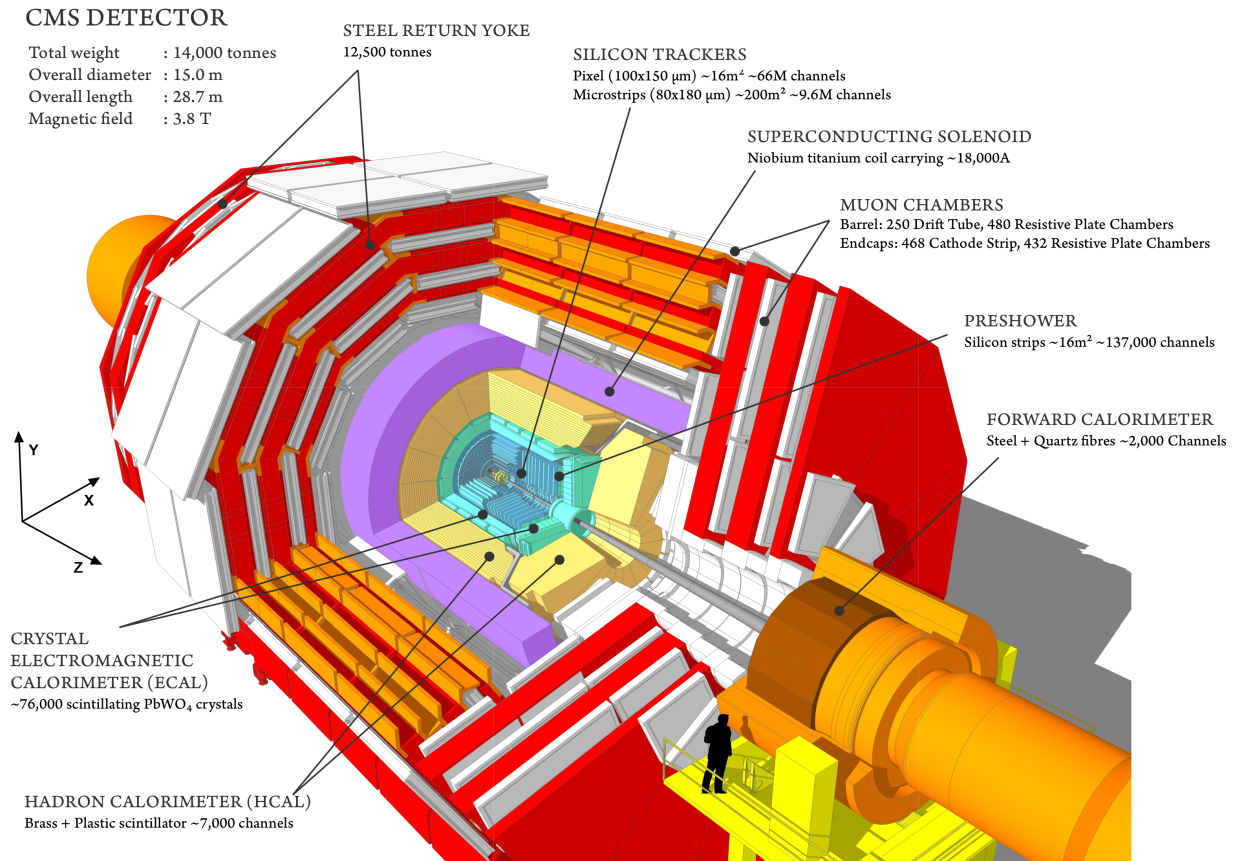


FIGURE 1.6 – Vue schématique du détecteur CMS.

CMS répond aux exigences des conditions de fonctionnement du LHC, qui lui imposent une grande granularité afin d'être capable de traiter le nombre d'événements superposés lors d'un croisement de faisceaux (*Pile-Up*), ainsi qu'une bonne résolution temporelle (avec nominale un croisement de faisceaux toutes les 25 ns) et la résistance aux radiations. De plus, il a été conçu dans le souci d'une herméticité maximale, de façon à détecter la quasi-totalité des particules issues des collisions.

Le système de coordonnées utilisé dans CMS a pour origine le point de collision nominal, l'axe y vertical dirigé vers le haut, l'axe x horizontal pointant vers le centre du LHC, l'axe z dans la direction des faisceaux de façon à former un trièdre direct, illustré dans la Fig. 1.6. L'angle azimutal ϕ est mesuré à partir de l'axe x dans le plan xy , l'angle polaire θ est mesuré à partir de l'axe z , la distance radiale transverse r_T est définie par la distance à l'origine de la projection dans le plan transverse xy . La pseudorapidité η , très utilisée en physique des particules sur accélérateur car elle présente l'avantage d'une quantité de particules produites constante par unité, est définie par $\eta = -\ln(\tan\theta/2)$.

La collaboration autour de ce détecteur, initiée en 1990, atteint aujourd'hui plus de 3800 scientifiques répartis dans 42 pays.

1.2.3 Le détecteur de traces (trajectographe)

Détecteur d'importance primordiale pour la reconstruction des électrons, le trajectographe est dédié à la mesure de la trajectoire des particules chargées issues des collisions, pour une large gamme en impulsion transverse p_T ($100 \text{ MeV} < p_T \lesssim 1 \text{ TeV}$), et conséquemment à la reconstruction des points de collision (vertex) primaires et secondaires des protons de chaque événement. Détecteur cylindrique de 5.8m de longueur et 2.5 m de diamètre situé au plus près de la collision, il est entièrement constitué de silicium (200 m^2 au total) et offre une couverture jusqu'à $|\eta| \leq 2.5$. Deux parties le composent : d'une part le détecteur le plus proche du point d'interaction avec une très fine granularité (pixels) et d'autre part le plus externe en micro-pistes. Sa structure est visible sur la Fig. 1.7 où il est représenté en coupe longitudinale, et en coupe transverse pour la partie centrale.

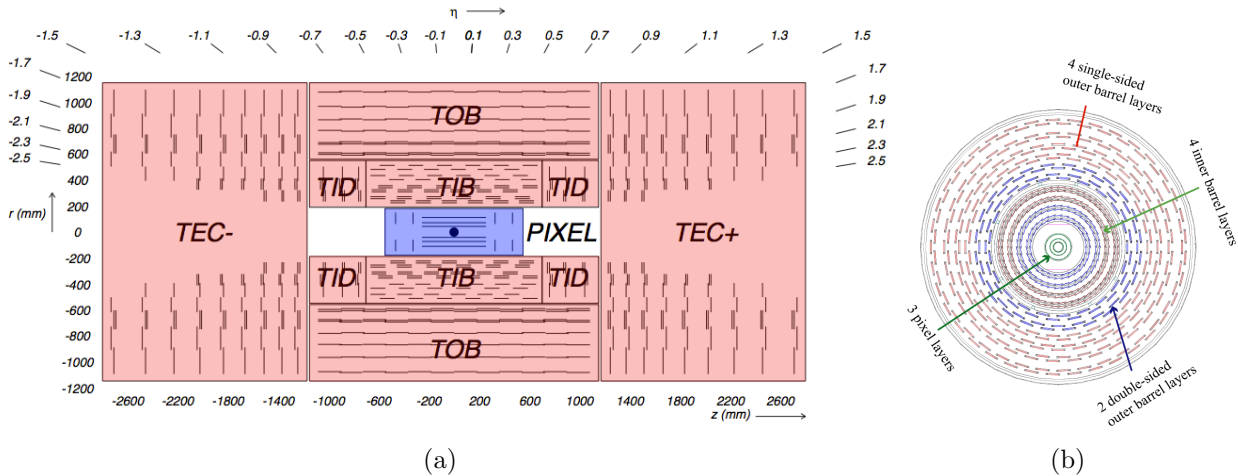


FIGURE 1.7 – Vue schématique du trajectographe de CMS. (a) Vue en coupe longitudinale. (b) Vue en coupe transverse de la partie centrale.

Le détecteur à pixels est constitué de 3 couches centrales, à distances radiales transverses $r_T = 4.4, 7.3$ et 10.2 cm couvrant une longueur $|z| \leq 53$ cm, complétées à l'avant par des bouchons constitués chacun de 2 disques placés à $|z| = 34.5$ et 46.5 cm et couvrant $6 \leq r_T \leq 15$ cm. Il est représenté dans la Fig. 1.8 (a). Chaque pixel mesure $100 \times 150 \mu\text{m}$. Ce détecteur au plus proche de la collision représente 66 millions de canaux. Sa géométrie lui permet de fournir 3 points de mesure par trace sur quasiment toute la gamme $|\eta| < 2.5$.

Le détecteur à micro-pistes est constituée de 10 couches centrales - TIB (*Tracker Inner Barrel*) et TOB (*Tracker Outer Barrel*) dans la Fig. 1.7 - complétées par 12 disques de chaque côté - TID (*Tracker Inner Disks*) et TEC (*Tracker EndCaps*) dans la Fig. 1.7. Une photo du TIB lors de sa construction est représentée dans la Fig. 1.8 (b). Les couches centrales sont situées à distances radiales transverses comprises entre 25 et 108 cm et couvrent une longueur $|z| \leq 70$ cm pour la partie la plus interne, $|z| \leq 109$ cm pour la partie la plus externe. Les 3 disques internes couvrent $80 \leq |z| \leq 90$ cm et $20 \leq r_T \leq 50$ cm et les 9 externes couvrent $124 \leq |z| \leq 280$ cm et $r_{\min} \leq r_T \leq 113$ cm, r_{\min} variant de 23 à 31 cm. Ce détecteur représente 9.3 millions de canaux et permet de fournir jusqu'à 14 (17) points de mesure par trace dans la partie centrale (à l'avant).

Pour des muons isolés émis dans la région centrale ($|\eta| < 1.4$) avec $p_T = 100$ GeV, les résolutions obtenues sont $\approx 2.8\%$ en p_T , et ≈ 10 et $30 \mu\text{m}$ en paramètres d'impact transverse et longitudinal respectivement. Pour les mêmes muons avec $p_T = 10$ GeV, les résolutions sont respectivement 0.9% , 20 et $35 \mu\text{m}$. La résolution en position pour les vertex d'interactions reconstruits est $\approx 10 - 12 \mu\text{m}$ pour chacune des trois dimensions [23].

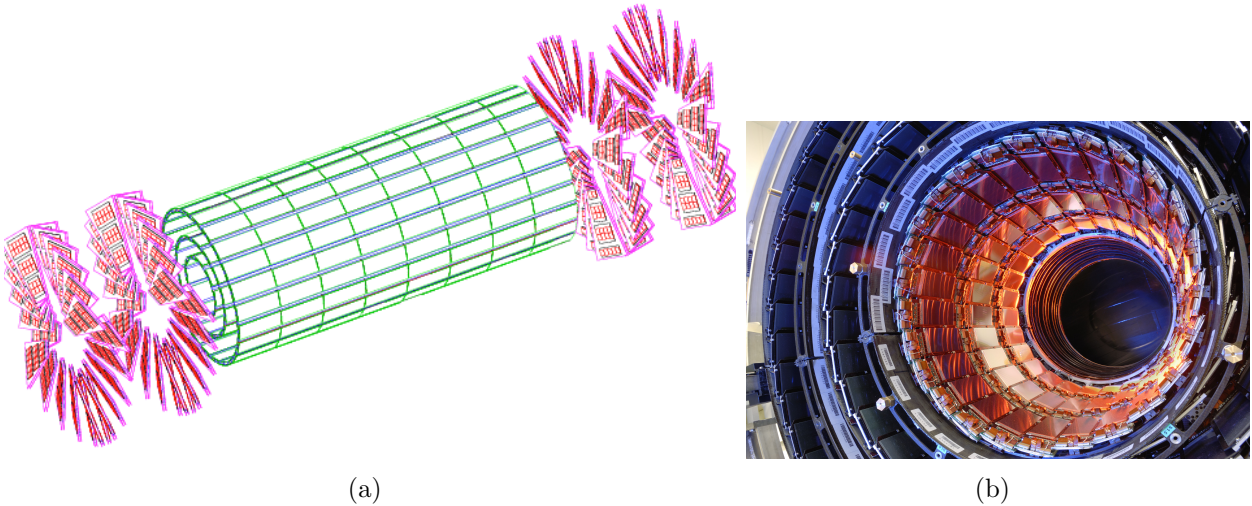


FIGURE 1.8 – (a) Schéma du trajectographe à pixels. (b) Photo du TIB en construction.

La constitution du trajectographe implique une quantité de matière importante présente devant le calorimètre électromagnétique, principalement à cause de la structure mécanique, de système de refroidissement et de services. Celle-ci est représentée dans la Fig. 1.9 en fonction de η . Quantifiée en longueur relative de radiation X/X_0 ¹, elle s'accroît de 0.4 à $|\eta| \approx 0$ à 1.8 à $|\eta| \approx 1.4$, puis diminue jusqu'à 1 à $|\eta| \approx 2.5$. En particulier cette quantité de matière, traversée par les électrons avant d'atteindre le ECAL, induit une perte non négligeable d'énergie par bremsstrahlung, ainsi que la conversion de ces photons de bremsstrahlung. Cet effet, combiné avec la présence du champ magnétique intense, implique des répercussions conséquentes sur la reconstruction des électrons.

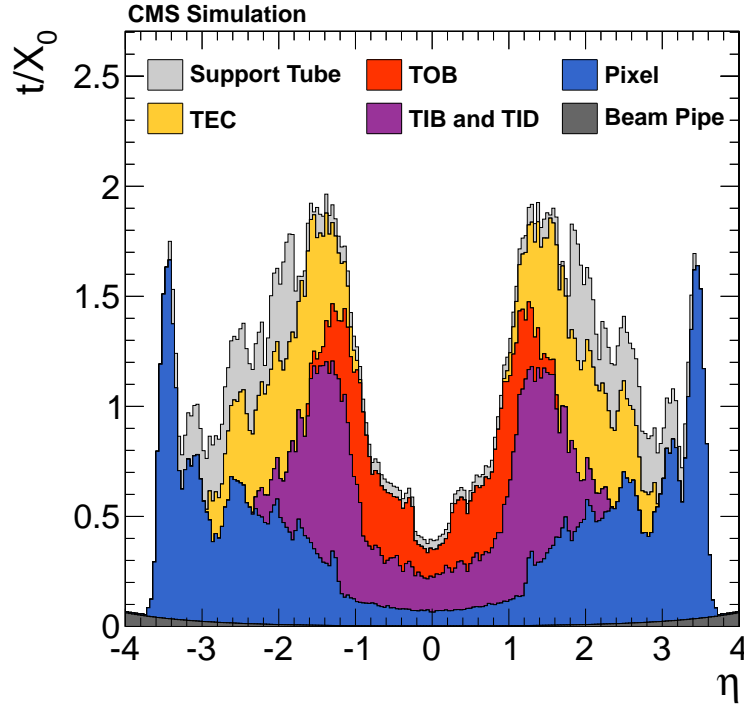


FIGURE 1.9 – Quantité de matière du trajectographe en fonction de η , exprimée en longueurs relatives de radiation X/X_0 des différentes composantes du trajectographe en fonction de η .

1. X_0 est la distance moyenne pour laquelle un électron de haute énergie ne conserve que 1/e de son énergie, le reste étant perdu par bremsstrahlung.

1.2.4 Le calorimètre électromagnétique (ECAL)

1.2.4.1 Détection

Responsable de la mesure d'énergie des électrons et photons qu'il arrête, le ECAL est un calorimètre homogène et hermétique, constitué de cristaux scintillants de Tungstate de Plomb ($PbWO_4$). Il est composé d'une partie cylindrique centrale (*barrel*) couvrant $|\eta| \leq 1.479$ dont la surface interne est située à $r_T = 129$ cm, complétée par deux bouchons (*endcaps*) couvrant $1.479 \leq |\eta| \leq 3$, placés à $|z| = 315.4$ cm. Ces derniers sont dotés sur leur face interne de détecteurs de pieds de gerbe (*preshowers*), de 20 cm d'épaisseur couvrant $1.653 \leq |\eta| \leq 2.6$. Les *preshowers* sont constitués de deux plans de détecteurs en micro-pistes de silicium (avec une orientation perpendiculaire entre les plans), chacun précédé d'une couche de plomb permettant l'amorce de la gerbe électromagnétique. Ils représentent un total de $3 X_0$, ainsi $\approx 5\%$ de l'énergie d'un photon ou d'un électron émis à l'avant sont déposés dans le *preshower*. La géométrie du ECAL est illustrée dans les Figs. 1.10 (a) en représentation 3D, et 1.11 en coupe longitudinale.

Le Tungstate de Plomb constituant les cristaux a été choisi pour son importante densité (8.28 g.cm^{-3}), sa faible longueur de radiation ($X_0 = 0.89 \text{ cm}$) et son faible rayon de Molière (2.3 cm)², autorisant ainsi une conception du ECAL compacte et une granularité fine. La résistance aux radiations de ces cristaux, ainsi que leur rapidité de réponse (80 % de la lumière est émise en moins de 25 ns) permet au ECAL de fonctionner efficacement dans l'environnement imposé par le LHC. Les pertes de transparence subies au cours du temps du fait des radiations sont contrôlées et corrigées grâce à l'injection de lumière laser de différentes longueurs d'onde dans les cristaux. Une photo d'un des derniers cristaux des *endcaps* et la simulation d'une gerbe d'électron dans un cristal sont montrées dans les Figs. 1.10 (b) et 1.10 (c).

Le *barrel* est constitué de 61200 cristaux trapézoïdaux, de section droite $22 \times 22 \text{ mm}^2$ pour la face avant (entrée des particules dans le calorimètre) soit une granularité de $0.0174 \times 0.0174 \text{ rad}$ en $\eta \times \phi$, avec une longueur de 230 mm soit $25.8 X_0$. Quatre modules, contenant chacun 4×5 cristaux en η (5×5 pour le plus proche de $|\eta| = 0$) et 4×5 cristaux en ϕ , forment un supermodule de 1700 cristaux couvrant 20 degrés en ϕ sur la moitié de la longueur totale, de sorte que 18 supermodules de chaque côté de l'axe $\eta = 0$ couvrent l'ensemble du *barrel*. Les cristaux sont agencés en une géométrie quasi-projective, décalés de 3 degrés en η et ϕ par rapport à l'axe projectif passant par 0, afin d'éviter que les électrons ou photons ne s'insèrent dans des zones non instrumentées entre les cristaux, les modules ou les supermodules (*cracks*).

Les *endcaps* sont constitués de 14648 cristaux également trapézoïdaux (7324 dans chaque *endcap*), de section droite $28.62 \times 28.62 \text{ mm}^2$ pour la face avant, avec une longueur de 220 mm soit $24.7 X_0$. Les cristaux sont assemblés en matrices de 5×5 cristaux formant chacune un su-

2. Le rayon de Molière caractérise l'extension latérale de la gerbe électromagnétique : il est défini par le rayon d'un cylindre contenant en moyenne 90 % de l'énergie déposée par la gerbe.

percristal. Chaque *endcap* est séparé en deux demi-disques (*Dees*). Les cristaux sont agencés en une géométrie quasi-projective, pointant 1300 mm avant le point 0, impliquant ainsi des décalages de 2 à 8 degrés par rapport à l'axe projectif passant par 0.

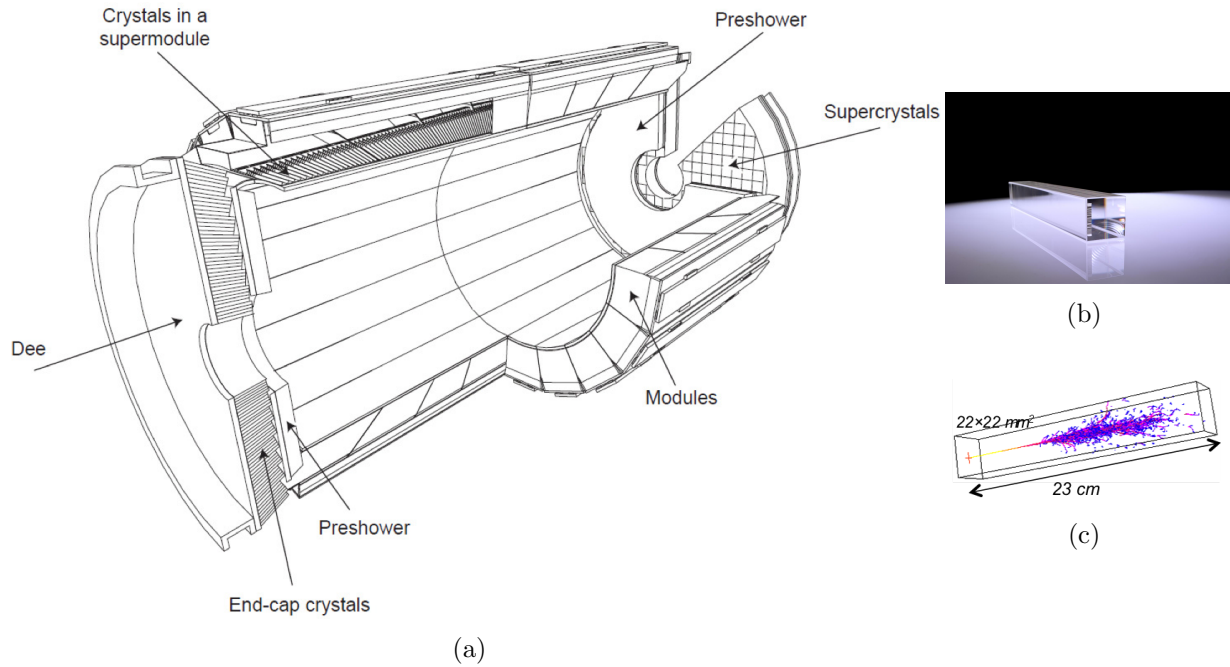


FIGURE 1.10 – (a) Vue éclatée du ECAL. (b) Photo d'un des derniers cristaux des *endcaps*. (c) Schéma d'une gerbe d'un électron dans un cristal.

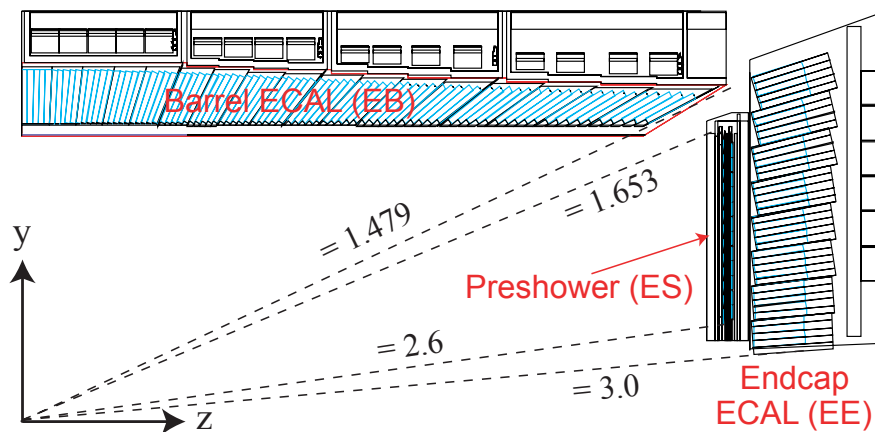


FIGURE 1.11 – Schéma du ECAL en coupe transverse.

1.2.4.2 Électronique de lecture

Les cristaux sont équipés de photodétecteurs recueillant la lumière émise. Ceux-ci doivent être rapides, résistants aux radiations et capables de fonctionner dans le champ magnétique de CMS. De plus, la lumière émise par les cristaux étant relativement faible, ils jouent un rôle amplificateur, tout en étant le plus insensible possible aux autres particules les traversant (une sensibilité à des neutrons par exemple a pour effet l'illusion de dépôts d'énergie considérables localisés dans un unique cristal). La différence du taux d'irradiation, plus intense à l'avant, a déterminé un choix différent entre le *barrel* et les *endcaps* : des photodiodes à avalanche (*Avalanche PhotoDiode*, APD) pour les cristaux du *barrel* et des phototriodes à vide (*Vacuum PhotoTriode*, VPT) dans les *endcaps*. Deux APD, chacune d'une surface active de $5 \times 5 \text{ mm}^2$, sont fixées par cristal du *barrel* et fonctionnent en gain 50, à une température de 18° . Les VPT des cristaux des *endcaps* ont une surface active de 280 mm^2 . Plus résistantes aux radiations et moins sensibles au champ magnétique, elles ont un gain effectif plus faible, compensé par leur surface de collection plus importante.

Les signaux issus des photo-détecteurs doivent être lus de façon rapide et précise. Le système électronique de lecture, schématisé dans la Fig. 1.12, est composé de l'électronique embarquée sur le détecteur et de l'électronique hors détecteur.

Le système électronique installé sur le détecteur lit chaque tour de *trigger* (5×5 cristaux en $\eta \times \phi$) dans le *barrel*, et chaque supercristal (5×5 cristaux en $x \times y$) dans les *endcaps*. Pour chaque tour de *trigger* (ou supercristal) il est constitué principalement de 5 cartes *Very Front End* (VFE), d'une carte *Front End* (FE), et d'une carte mère. Dans les VFE la lumière collectée dans cinq photodétecteurs est mise en forme et amplifiée, puis digitisée à la fréquence des croisements de faisceaux et enfin sommée. La carte FE stocke les données digitisées de 5 cartes VFE pendant le temps de latence du système de déclenchement de niveau 1 ($3.2 \mu\text{s}$). De plus, elle initie la génération des primitives de déclenchement (*Trigger Primitive*, TP), briques élémentaires contenant les informations sur les tours de *trigger* pour le déclenchement, en sommant les énergies issues des VFE (5 cristaux en ϕ par VFE) et pour le *barrel* en calculant la compacité du signal dans la tour de *trigger*, utile pour identifier une signature d'électron ou de photon.

Le système électronique hors détecteur chemine les informations vers l'acquisition et le déclenchement. Dans la voie de déclenchement, les *Trigger Concentrator Cards* (TCC) finalisent et synchronisent les TP, commencées dans les FE, avant transmission pour chaque croisement de faisceaux au *trigger* calorimétrique régional, sous-système du déclenchement de niveau 1. Une TP contient, pour chaque tour de *trigger*, 8 bits pour l'énergie transverse de la tour et un bit de *Fine Grain* (FG) représentant la compacité du dépôt d'énergie dans la tour. Une carte TCC reçoit les informations de 68 tours de déclenchement dans le *barrel*, et un maximum de 48 entrées dans les *endcaps*. Si un signal de déclenchement est envoyé par le niveau 1, une classification des tours est effectuée par les TCC en fonction de leur énergie transverse : niveau d'intérêt élevé si $E_T > 2 \text{ GeV}$, intermédiaire si $1 < E_T < 2 \text{ GeV}$ et faible

si $E_T < 1 \text{ GeV}$.

Cette classification est envoyée au *Selective Readout Processor* (SRP) [24] qui détermine le mode de lecture pour chaque tour : lecture complète des cristaux de la matrice 3×3 des tours centrées sur une tour d'intérêt élevé, uniquement des cristaux de la tour si celle-ci est jugée d'intérêt intermédiaire, et des cristaux de la tour avec “suppression de zéro” si celle-ci est d'intérêt faible. La “suppression de zéro” supprime la lecture du cristal si son énergie transverse est plus faible que 96 MeV (*barrel*) ou 360 MeV (*endcaps*), représentant entre 2 et 3 écart-types de la distribution du bruit dans un cristal. Cette opération permet de réduire le flot de données du ECAL du facteur ≈ 20 nécessaire à une lecture fonctionnelle.

Dans la voie d'acquisition, les données sont envoyées des FE vers les *Data Concentrator Cards* (DCC) qui récupèrent également les sorties du SRP et des TCC. Ensuite, ces cartes produisent le filtrage des données de chaque cristal basé sur les informations du SRP. Après ce filtrage et une mise en forme, les informations de l'événement sont envoyées au système central d'acquisition de CMS.

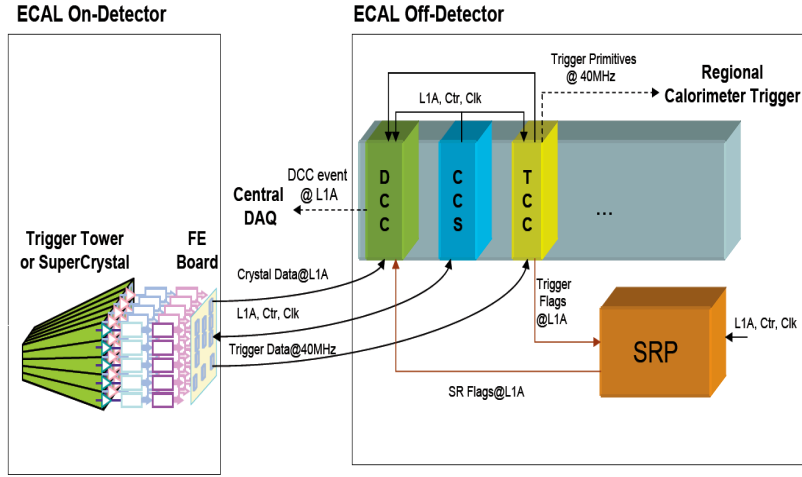


FIGURE 1.12 – Architecture de l'électronique de lecture du ECAL [24].

1.2.4.3 Résolution en énergie

La résolution intrinsèque du *barrel* a été mesurée en tests faisceaux, sous une incidence normale pointant le centre d'un cristal à $\pm 4 \times 4 \text{ mm}^2$, autour duquel une matrice de 3×3 cristaux est utilisée pour mesurer l'énergie de l'électron incident [25]. Dans ces conditions la résolution s'exprime de la façon suivante :

$$\frac{\sigma_E}{E} = \frac{2.8\%}{\sqrt{E}} \oplus \frac{12\%}{E} \oplus 0.3\% \quad (1.1)$$

Le terme stochastique de $\frac{2.8\%}{\sqrt{E}}$ est relié à la fluctuation dans la détection des particules de la gerbe (au niveau de la collection de la gerbe, du nombre de photo-électrons émis et de la variation du gain). Le terme de $\frac{12\%}{E}$ correspond au bruit électronique de 40 MeV par cristal (auquel doit s'ajouter en conditions de collisions le bruit résultant de l'empilement des événements). Le terme constant de 0.3%, dominant à haute énergie, prend en compte les effets induisant des variations de la réponse du détecteur en fonction de différents paramètres (non-uniformité de la collection longitudinale, perte d'énergie à l'arrière, précision d'intercalibration des cristaux).

1.2.5 Le calorimètre hadronique (HCAL)

Le HCAL est un calorimètre à échantillonnage constitué de couches de laiton et de scintillateur plastique, couvrant $|\eta| < 2.9$, avec une granularité en $\eta \times \phi$ de 0.087×0.087 rad dans le *barrel*, et de 0.17×0.17 rad dans les *endcaps*, la granularité exacte dépendant de $|\eta|$. A l'avant la couverture est étendue avec un calorimètre hadronique en fer et fibre de quartz couvrant jusqu'à $|\eta| < 5.2$.

Ce premier chapitre, dédié à la production des bosons de Higgs, du point de vue des cadres théorique et expérimental, a décrit les premiers ingrédients nécessaires à la réalisation du sujet de ce mémoire. À ce stade, nous sommes donc pourvus de bosons de Higgs, produits grâce aux collisions des protons au LHC, nous connaissons dans quel canal de désintégration les rechercher et les détecteurs correspondants. Le prochain chapitre est naturellement consacré à une étape cruciale dans notre cheminement, l'enregistrement effectif de ces événements si soigneusement choisis.

Chapitre 2

Déclencher

Sommaire

2.1	Système de déclenchement	26
2.2	Déclenchement pour les électrons	27
2.2.1	Niveau 1 (L1)	27
2.2.1.1	Chaîne de déclenchement	27
2.2.1.2	Réjection des signaux anormaux	29
2.2.2	Niveau supérieur (HLT)	30
2.3	Déclenchement pour l'analyse du boson de Higgs en quatre leptons	31
2.3.1	Sélection des événements Higgs en quatre leptons	31
2.3.2	Sélection des événements pour les mesures de performance des électrons	34
2.4	Études d'efficacité du <i>trigger</i> diélectron	35
2.4.1	Efficacités par électron	35
2.4.2	Efficacités du <i>trigger</i> diélectron pour les événements Higgs en quatre électrons	39

Maintenant que nous savons comment sont produits et à quoi ressemblent les bosons de Higgs qui nous intéressent, la prochaine étape consiste à s'assurer que ceux-ci soient bien enregistrés par CMS. L'opération est complexe. D'une part, comme indiqué dans le chapitre précédent, ≈ 9 ordres de grandeurs séparent la section efficace totale de celle de boson de Higgs. D'autre part, la sélection est rude pour les événements qui ont pour ambition d'accéder à une analyse : d'un taux de croisement de faisceaux nominal à 40 MHz (avec plusieurs collisions simultanées par croisement de faisceaux, une vingtaine à la luminosité nominale), seuls ≈ 400 Hz sont enregistrés par CMS, soit une réduction drastique d'au moins 5 ordres de grandeur. Enfin, le programme de physique de CMS est loin de se limiter à la recherche du boson de Higgs... Lors du tri en ligne des événements issus des collisions, il s'agit d'identifier efficacement lesquels sont susceptibles de ressembler à des bosons de Higgs afin de les enregistrer, sachant que les événements non sélectionnés sont simplement perdus. Ce chapitre décrit le déclenchement des événements du point de vue des électrons, et en particulier pour l'analyse du boson de Higgs se désintégrant en quatre leptons. Il y est fait fréquemment appel à la reconstruction et la sélection des électrons dont les détails, qui ne sont pas nécessaires pour la compréhension de ce chapitre, seront amplement exposés dans le chapitre suivant.

2.1 Système de déclenchement

Le système de déclenchement (*trigger*) de CMS est séparé en deux niveaux distincts : le premier (*Level1*, L1) est basé sur l'électronique programmable de lecture au niveau des détecteurs et aux alentours, et permet de réduire le taux d'un facteur 400, de 40 MHz à 100 kHz ; le second (*High Level Trigger*, HLT) est constitué d'une ferme de processeurs diminuant le taux d'un facteur 250, de 100 kHz à ≈ 400 Hz acquis par CMS [22].

Le L1 utilise les informations de basse granularité provenant des calorimètres et des détecteurs à muons, afin de sélectionner les événements à l'aide de critères simples [26, 27], avec une décision lors de chaque croisement de faisceaux. Pendant les $3.6 \mu\text{s}$ nécessaires à la prise de décision, la totalité des données est conservée en mémoire tampon. La génération des Primitives de déclenchement (*Trigger Primitive*, TP) constitue la première étape de l'algorithme du L1. Les TP sont élaborées à partir des dépôts d'énergie dans les tours de *trigger* des calorimètres et dans les segments de traces ou patterns de *hits* dans les chambres à muons. Les *triggers* régionaux combinent les informations des TP afin de déterminer des objets de *trigger* tels que des candidats électrons ou photons, jets, taus et muons dans des zones spatialement limitées. Ces objets sont classés et ordonnés en fonction de leur énergie ou impulsion transverses et de leur qualité. Les *triggers* globaux, calorimétrique d'une part et pour les chambres à muons d'autre part, déterminent les objets les mieux classés pour le détecteur dans son ensemble, ainsi que des observables globales telles que l'énergie transverse manquante, et les transmettent au *trigger* général qui prend la décision de rejeter un événement ou de l'accepter pour le transmettre au HLT.

Le HLT accède à la totalité des données lues par l'ensemble des sous-détecteurs et effectue la reconstruction des objets et observables de façon similaire à la reconstruction hors ligne, en combinant les informations de plusieurs détecteurs [28]. Le système, en constante évolution durant la prise de données, est typiquement composé de 13000 cœurs de processeurs, ce qui représente 20000 processeurs d'événements grâce à la technique d'*Hyper-Threading* consistant à créer deux processeurs logiques sur une seule puce (appliquée à la majorité des cœurs). Le temps de décision en 2012 était ainsi de 165 ms par événement [29]. La stratégie des algorithmes du HLT est de s'appuyer sur les objets et informations issus du L1 d'une part, et d'autre part de séparer la sélection des événements en différentes étapes successives avec une reconstruction partielle et une sélection à chaque niveau, permettant ainsi une réjection rapide des événements. Les algorithmes de déclenchement forment un menu, optimisé en fonction des conditions de prise de données et du programme d'étude de CMS. Si un événement est jugé intéressant par l'un des algorithmes du HLT, il est enregistré et continue son chemin vers l'analyse.

2.2 Déclenchement pour les électrons

2.2.1 Niveau 1 (L1)

2.2.1.1 Chaîne de déclenchement

La génération des TP pour le ECAL a été décrite dans le chapitre précédent, section 1.2.4.2. Celles-ci contiennent, pour chaque tour de *trigger* et pour chaque croisement de faisceaux, les informations de l'énergie transverse et de la compacité en η du dépôt d'énergie (*Fine Grain*, FG). Le FG détermine si deux bandes adjacentes en η (de 5 cristaux chacune) contiennent une fraction significative de l'énergie transverse de la tour (typiquement 90 %).

Les TP sont transmises au *trigger* calorimétrique régional (*Regional Calorimeter Trigger*, RCT). Les candidats électrons ou photons (sans distinction entre eux) sont construits à l'aide d'un algorithme de fenêtre glissante de 3×3 tours de *trigger* en $\eta \times \phi$ s'appliquant dans la totalité du ECAL, illustré dans la Fig. 2.1 [30] : à l'intérieur de la fenêtre, si la tour centrale possède une énergie transverse maximale par rapport à ses plus proches voisines, elle est sélectionnée (Hit) pour former un candidat électron/photon (EG) avec la tour d'énergie transverse maximale parmi ses quatre plus proches voisines (Max). L'énergie du candidat EG est la somme des énergies des deux tours de *trigger* formant le candidat. Des critères d'identification sont alors appliqués dépendant de l'énergie transverse des candidats. Les électrons ou photons produisant des cascades électromagnétiques dans le ECAL caractérisées par une extension en η compacte, la tour centrale des candidats doit vérifier le critère de compacité déterminé par le FG si son énergie transverse E_T excède 6 GeV. Les cascades électromagnétiques étant essentiellement contenues dans le ECAL, la tour centrale des candidats doit vérifier $H_T/E_T < 5\%$ si $E_T > 2$ GeV, H_T étant l'énergie transverse déposée dans la tour de *trigger* HCAL à l'arrière de la tour ECAL considérée. De plus, pour qu'un

candidat soit considéré comme isolé, deux critères additionnels sont appliqués. D’une part, les huit tours voisines de la tour centrale doivent vérifier les deux critères précédents (FG et H_T/E_T). D’autre part, l’énergie transverse totale d’au moins l’un des coins formés par 5 tours adjacentes directement autour de la tour centrale doit vérifier $E_T^{\text{coin}} < 3.5 \text{ GeV}$.

Le *trigger* calorimétrique global (*Global Calorimeter Trigger*, GCT) reçoit 72 candidats EG de la catégorie des non-isolés et 72 de la catégorie des isolés, correspondant pour chaque catégorie aux 4 candidats de plus haute énergie transverse par secteur couvrant la moitié du ECAL en η et 40° en ϕ [31]. Le GCT en sélectionne les 4 candidats de plus haute énergie transverse pour chaque catégorie et les envoie au *trigger* global (*Global Trigger*, GT). Le GT est responsable de la génération de la décision finale du L1 d’accepter ou de rejeter l’événement, en fonction des algorithmes de sélection du menu de *trigger*. Un maximum de 128 algorithmes différents peuvent être implémentés dans un menu de L1. En ce qui concerne les événements contenant des électrons, ils sont sélectionnés avec la présence d’un ou plusieurs candidats EG, parfois associés avec des candidats d’autres types, comme les muons par exemple. Différents critères en E_T , isolation, localisation, sont appliqués aux candidats EG.

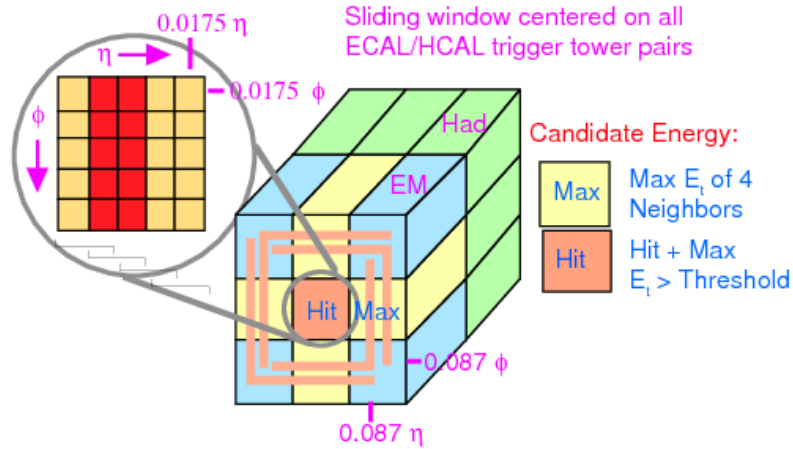


FIGURE 2.1 – Algorithme de déclenchement du L1 pour les électrons/photons. Un candidat L1 est formé par une tour centrale d’énergie transverse maximale locale (Hit) dans une fenêtre glissante en 3×3 tours de *trigger* et par la tour de plus haute énergie transverse parmi ses 4 plus proches voisines (Max). La compacité en η du dépôt d’énergie (FG) est illustrée dans la partie zoomée. Les tours de *trigger* HCAL sont montrées à l’arrière des tours ECAL. Les coins formés par 5 tours adjacentes autour de la tour centrale, utilisés pour l’isolation des candidats, sont montrés par les lignes orange [30].

2.2.1.2 Réjection des signaux anormaux

Des signaux anormaux (*spikes*) ont été observés dès les premières collisions du LHC dans la partie *barrel* du ECAL [32]. Un *spike* est caractérisé par un signal isolé dans un cristal unique, de grande énergie apparente, et est engendré par une ionisation dans les APD reliée à la présence de neutrons. Le taux d'apparition de ce phénomène est proportionnel à l'intensité et à l'énergie des faisceaux, et à défaut d'être drastiquement éliminé en ligne, augmente indésirablement le taux de déclenchement du L1 jusqu'à saturer la bande passante à haute luminosité.

Lors de la reconstruction hors ligne, la stratégie de réjection repose sur deux critères, l'un topologique et l'autre temporel, discernant les *spikes* des signaux électromagnétiques réels. Une cascade électromagnétique est distribuée sur plusieurs cristaux, avec jusqu'à 80 % de l'énergie incidente déposée dans le cristal central et la grande majorité de l'énergie restante dans les 4 cristaux adjacents. Ainsi le premier critère restreint l'observable C_s , dénommée "croix suisse" et définie comme $C_s = 1 - E_4/E_1$, avec E_1 l'énergie du cristal central et E_4 la somme des énergies des 4 cristaux adjacents, à ne pas dépasser 95 % (100 % étant la valeur pour laquelle le cristal central est entièrement isolé par rapport à des plus proches voisins), et permet de rejeter plus de 99 % des *spikes* d'énergie supérieure à 10 GeV avec une perte d'efficacité négligeable. D'autre part, la forme du signal en fonction du temps est différente pour les *spikes*, pour lesquels la constante de temps de scintillation des cristaux (10 ns) n'intervient pas car le signal est directement produit dans les APD. Ainsi le deuxième critère restreint le temps mesuré du signal à ± 3 ns du croisement de faisceaux à l'origine de l'événement considéré, et permet de rejeter plus de 90 % des *spikes* avec un impact négligeable sur l'efficacité (la résolution en temps pour les signaux électromagnétiques étant inférieure à 1 ns).

Pour le déclenchement en ligne, une observable topologique, semblable à C_s mais unidimensionnelle, a été développée, utilisant le *Strip Fine Grain* (sFG) qui identifie une tour de *trigger* correspondant à une cascade électromagnétique dès lors qu'au moins une de ses bandes en $\eta \times \phi = 1 \times 5$ contient deux cristaux avec une énergie transverse supérieure à un seuil E_T^{seuil1} , programmable et de l'ordre de quelques centaines de MeV. Ainsi, si une tour ne passe pas ce critère et que son énergie transverse est supérieure à un second seuil E_T^{seuil2} , elle est considérée comme un *spike* et ne contribue pas au déclenchement pour l'événement considéré. Les seuils E_T^{seuil1} et E_T^{seuil2} ont été optimisés en fonction des périodes de prise de données, afin de conserver une excellente efficacité de déclenchement pour les électrons et les photons tout en rejetant le maximum de *spikes*, et sont montrés dans la Table 2.1 pour les années 2011 et 2012. Les performances de l'algorithme sont similaires pour les deux années, montrant une réduction d'un facteur > 4 du taux de déclenchement pour la présence d'un candidat EG avec $E_T > 15$ GeV (et correspondant à 96 % des *spikes* rejetés dans les conditions de prise de données en 2010) et une efficacité $\gtrsim 99$ % pour les électrons avec $p_T > 15$ GeV.

Année	E_T^{seuil1} (MeV)	E_T^{seuil2} (GeV)
2011	258	8
2012	350	12

TABLE 2.1 – Valeurs des deux seuils utilisés dans l’identification en ligne des *spikes* pour les périodes de prise de données de 2011 et 2012.

2.2.2 Niveau supérieur (HLT)

Les algorithmes de sélection des candidats électrons au HLT [27] sont initiés par les candidats EG du L1. Chaque *trigger* du HLT est basé sur un algorithme de L1 déterminé, ainsi les événements étudiés doivent en premier lieu être sélectionnés par cet algorithme de L1. Puis une succession d’étapes de reconstruction et de sélection est réalisée, de façon à rejeter au plus tôt les événements qui ne seront pas sélectionnés.

La première étape est purement calorimétrique et reconstruit un agrégat d’énergie, appelé *supercluster* (SC), autour des candidats EG du L1 en regroupant les cristaux de la même façon que la reconstruction hors ligne (qui sera décrite en détail dans le chapitre 3) afin de collecter efficacement l’énergie déposée par un électron dans le ECAL. Un premier ensemble de sélections est appliqué, basé sur l’énergie transverse (E_T) des candidats, sur les propriétés de la cascade électromagnétique, à l’aide de l’observable de forme $\sigma_{\eta\eta}$ qui décrit l’extension de la cascade en η (détaillée dans la Section 3.5.1), puis sur l’isolation dans le ECAL ($\text{Iso}_{\text{ECAL}}/E_T$), qui limite l’activité relative autour du candidat dans un cône de taille $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.3$ en excluant le candidat par des critères topologiques.

L’information provenant du HCAL est alors exploitée pour les candidats vérifiant les sélections précédentes. Un premier critère limite la quantité d’énergie présente à l’arrière du SC, à l’aide de l’observable H/E_{SC} , E_{SC} étant l’énergie du SC et H la somme des énergies dans les tours du HCAL à l’intérieur d’un cône de taille $\Delta R = 0.15$ autour de la direction du SC. Le deuxième limite l’activité relative dans le HCAL dans un cône de taille $\Delta R = 0.3$ autour du SC ($\text{Iso}_{\text{HCAL}}/E_T$).

La dernière étape reconstruit une trace à l’aide d’un filtre de Kalman (KF) [23], dont les deux premiers *hits* sont compatibles avec la direction du SC. Une fois la trace reconstruite, des critères de compatibilité avec le SC sont appliqués, cinématiques et/ou géométriques, à l’aide de la différence de direction en η ($\Delta\eta$) ou en ϕ ($\Delta\phi$) entre le SC et la trace. Un critère d’isolation concernant la somme relative des impulsions transverses des traces dans un cône de taille $\Delta R = 0.3$ autour de la trace du candidat ($\text{Iso}_{\text{TR}}/E_T$) est également appliqué.

Dans le cas où l’algorithme requiert deux candidats (ou plus) avec des sélections différentes, les conditions sur le deuxième candidat sont appliquées si le premier candidat est sélectionné (et ainsi de suite). Enfin, les critères liant plusieurs candidats entre eux sont considérés, en particulier la différence entre les positions z du point d’approche minimale de la collision pour chacune des traces (δz) et/ou la masse invariante entre deux candidats.

2.3 Déclenchement pour l'analyse du boson de Higgs en quatre leptons

2.3.1 Sélection des événements Higgs en quatre leptons

Les algorithmes de *trigger* développés pour l'analyse du boson de Higgs en quatre leptons sélectionnent des événements contenant deux ou trois candidats leptons (électrons ou muons), avec des critères dépendant de leur type et de leur nombre [33, 34]¹. Ces algorithmes ont été optimisés dans le but d'obtenir une efficacité maximale de sélection pour ces événements rares sur toute la gamme de la masse du boson de Higgs, qui rappelons-le demeurait encore inconnue au moment de leur définition. Ceux-ci ont évolué en fonction des différentes périodes de prise de données afin de faire face à l'afflux croissant du nombre d'événements produits.

La Table 2.2 désigne les *triggers* utilisés en 2012 pour la sélection des événements de l'analyse du boson de Higgs en quatre leptons. En 2011, des *triggers* similaires ont été utilisés, toutefois avec des critères plus relâchés. Leur noms indiquent le nombre de candidats de chaque type, leurs seuils en E_T et leurs sélections, qui sont explicitées dans la Table 2.3 pour les critères sur les électrons [35]. L'efficacité de sélection obtenue grâce à la combinaison des algorithmes de *trigger* décrits précédemment est $> 99\%$ pour des événements de Higgs simulés avec une masse $m_H > 120$ GeV se désintégrant en quatre leptons et vérifiant les sélections finales de l'analyse.

Le *trigger* principal en électrons pour l'analyse du boson de Higgs en deux électrons et deux muons, ou quatre électrons, nommé `HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL`, requiert deux candidats avec des seuils en E_T égaux à 17 et 8 GeV et des sélections d'identification et d'isolation (nommés `Ele17` et `Ele8` par la suite), comme indiqué dans la Table 2.3, avec entre eux un $|\delta z| < 0.2$ cm. Pour le canal en deux électrons et deux muons, l'efficacité de déclenchement est complétée par l'utilisation d'un *trigger* dimuon ainsi que des *triggers* mixtes requérant un électron et un muon.

Dans le cas du canal en quatre électrons, l'efficacité est complétée par l'utilisation du *trigger* triélectron `HLT_Ele15_Ele8_Ele5_CaloIdL_TrkIdVL`. En effet, le *trigger* diélectron comprend des sélections sur les observables d'isolation, pour le calcul desquelles la présence d'autre leptons à l'intérieur du cône d'isolation n'est pas exclue, impliquant ainsi des pertes d'efficacité pour les événements contenant des leptons proches. À titre d'exemple, parmi les événements simulés de bosons de Higgs avec une masse de 120 GeV se désintégrant en quatre électrons,

1. En ce qui concerne les muons, les candidats associent une trace dans le trajectographe avec des *hits* dans les chambres à muons, avec une possibilité pour certains d'être reconstruits par la présence d'une trace dans le trajectographe et une collection partielle de *hits* dans les chambres à muons.

canal	HLT	L1
4e	HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL...	L1_DoubleEG_13_7
	..._Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL	
	OR HLT_Ele15_Ele8_Ele5_CaloIdL_TrkIdVL	L1_TripleEG_12_7_5
2e2 μ	HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL...	L1_DoubleEG_13_7
	..._Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL	
	OR HLT_Mu17_Mu8	L1_Mu10_MuOpen
	OR HLT_Mu17_TkMu8	L1_Mu10_MuOpen
	OR HLT_Mu17_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL	L1_Mu12_EG7
	OR HLT_Mu8_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL	L1_Mu3p5_EG12 OR L1_MuOpen_EG12
4 μ	HLT_Mu17_Mu8	L1_Mu10_MuOpen
	OR HLT_Mu17_TkMu8	L1_Mu10_MuOpen

TABLE 2.2 – Algorithmes de *triggers* utilisés pour la sélection des événements de l’analyse du boson de Higgs en quatre leptons. Sont indiqués pour chaque canal les algorithmes du HLT et du L1 à leur origine.

9 % présentent deux électrons dans un même cône d’isolation de taille $\Delta R = 0.3^2$. Le *trigger* triélectron a été développé spécialement afin de compenser cet effet, avec les sélections les plus relâchées possibles, et en particulier sans critère d’isolation. Il permet ainsi, pour une masse du boson de Higgs égale à 125 GeV, l’addition de 3.3 % d’efficacité pour atteindre 99 % dans le canal en quatre électrons.

La Table 2.4 montre les taux d’acceptation de l’ensemble des *triggers*, et en particulier des *triggers* di- et triélectron, pour l’un des derniers *runs* de prise de données de l’année 2012 (*Run* 208686, d’une durée de trois heures, enregistré le 05/12/2012) pour lequel l’intensité était maximale, avec une luminosité instantanée en début de *run* égale à $7.2 \times 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ et une luminosité totale délivrée égale à 55.1 pb^{-1} [36].

2. Pour 95 % de ces événements, les deux leptons proviennent de deux bosons Z différents.

Nom	Sévérité	Observable	Sélection <i>barrel</i>	Sélection <i>endcaps</i>
CaloId	VT (<i>Very Tight</i>)	$\sigma_{\eta\eta}$	< 0.011	< 0.031
		H/E_{SC}	< 0.05	< 0.05
	T (<i>Tight</i>)	$\sigma_{\eta\eta}$	< 0.011	< 0.031
		H/E_{SC}	< 0.15	< 0.1
	L (<i>Loose</i>)	$\sigma_{\eta\eta}$	< 0.014	< 0.035
		H/E_{SC}	< 0.15	< 0.10
CaloIso	VT (<i>Very Tight</i>)	$\text{ISO}_{\text{ECAL}}/E_{\text{T}}$	< 0.05	< 0.05
		$\text{ISO}_{\text{HCAL}}/E_{\text{T}}$	< 0.05	< 0.05
	VL (<i>Very Loose</i>)	$\text{ISO}_{\text{ECAL}}/E_{\text{T}}$	< 0.2	< 0.2
		$\text{ISO}_{\text{HCAL}}/E_{\text{T}}$	< 0.2	< 0.2
TrkId	T (<i>Tight</i>)	$ \Delta\eta $	< 0.008	< 0.008
		$ \Delta\phi $	< 0.07	< 0.05
	VL (<i>Very Loose</i>)	$ \Delta\eta $	< 0.01	< 0.01
		$ \Delta\phi $	< 0.15	< 0.10
TrkIso	VT (<i>Very Tight</i>)	$\text{ISO}_{\text{TR}}/E_{\text{T}}$	< 0.05	< 0.05
	VL (<i>Very Loose</i>)	$\text{ISO}_{\text{TR}}/E_{\text{T}}$	< 0.2	< 0.2

TABLE 2.3 – Sélections utilisées dans les algorithmes de *trigger* pour les électrons. Les observables concernent l’identification (Id) et l’isolation (Iso) dans les calorimètres (Calo) et le trajectographe (Trk). Le niveau des critères peut être sévère (*Tight*) ou relâché (*Loose*).

<i>Trigger</i>	Taux HLT (Hz)	Taux L1 (kHz)
Totalité	1093	64.3
diélectron	4	7.5
triélectron	3	2.6

TABLE 2.4 – Taux d’acceptation de la totalité des *triggers*, et des *triggers* di- et triélectron pour le *run* 208686. Le taux L1 maximal était de 84.4 kHz pour ce *run*.

2.3.2 Sélection des événements pour les mesures de performance des électrons

La sélection des événements entrant directement dans l'analyse principale constitue une étape nécessaire mais non suffisante. En effet, des événements complémentaires sont indispensables aux études inhérentes à l'analyse, en particulier la mesure des efficacités de sélection en ligne (*trigger*) et hors ligne (reconstruction, identification et isolation), du taux de mauvaise identification, et l'estimation du bruit de fond. Du point de vue des leptons, ces études exploitent principalement une source de leptons bien connue, les événements contenant des bosons Z se désintégrant en leptons, en utilisant la méthode dite de *tag-and-probe* (T&P) [37], qui sera décrite en détails dans la Section 3.6. En particulier en ce qui concerne les électrons, des événements $Z/\gamma^* \rightarrow e^+e^-$ sont sélectionnés, pour lesquels un électron vérifie des critères sévères et sert d'identifiant (*tag*), alors que le second électron est utilisé comme sonde (*probe*) pour la mesure considérée. Ces événements sont sélectionnés en ligne par les *triggers* de *tag-and-probe* décrits dans la Table 2.5. Leur noms indiquent le nombre de candidats de chaque type, leurs seuils en E_T et leurs sélections, qui sont explicitées dans la Table 2.3³. Il requièrent un électron avec des sélections sévères d'identification et d'isolation, et un second avec des sélections très relâchées (un électron reconstruit avec $E_T > 8$ GeV ou uniquement la présence d'un SC avec $E_T > 4$ GeV), avec une masse invariante des deux électrons supérieure à 50 GeV, permettant une sélection efficace des événements Z tout en rejetant significativement le bruit de fond. Ces *triggers*, présentant un taux important, subissent une réduction artificielle par un facteur de *prescale*, sélectionnant un événement tous les N_{prescale} événements acceptés par le *trigger*.

<i>Trigger</i> HLT	L1	$N_{\text{prescale}}^{\text{HLT}}$	HLT (Hz)	L1 (kHz)
HLT_Ele17_CaloIdVT_CaloIsoVT_TrkIdT_TrkIsoVT_...Ele8_Mass50	L1_DoubleEG_13_7	5	2	7.5
HLT_Ele20_CaloIdVT_CaloIsoVT_TrkIdT_TrkIsoVT_...SC4_Mass50	L1_SingleIsoEG18er	10	2	8.6

TABLE 2.5 – *Triggers tag-and-probe* pour les mesures concernant les électrons. Sont indiqués les *triggers* au HLT et au L1, leur facteur de *prescale* au HLT (pas de *prescale* au L1 pour ces *triggers*) et leurs taux après application de ce facteur pour le *run* 208686.

3. La dénomination “er” dans les *triggers* EG isolés de L1 signifie une restriction en η ($|\eta| < 2.172$) des candidats sélectionnés.

2.4 Études d'efficacité du *trigger* diélectron

Pour une analyse donnée, l'efficacité du déclenchement est en premier lieu estimée par la simulation, pour laquelle la description des algorithmes de *trigger* est très précise. Néanmoins il est important de mesurer cette efficacité avec les données, ou du moins de s'en approcher quand cela n'est pas directement possible, afin d'une part de vérifier que le comportement des *triggers* est bien celui attendu, et d'autre part d'estimer les systématiques engendrées par des différences entre la simulation et les données. La difficulté de cette mesure avec les données provient d'une part du fait que les événements enregistrés ont préalablement déclenché au moins un algorithme de *trigger*, et d'autre part, dans le cas particulier de l'analyse de boson de Higgs en quatre leptons, du fait qu'il n'existe pas de *trigger* sélectionnant le même type d'événements, mais suffisamment différent de celui considéré. Nonobstant, il est possible de réaliser certaines mesures utiles à l'estimation de l'efficacité du *trigger* diélectron pour les événements de données entrant dans l'analyse de boson de Higgs en quatre leptons, à l'aide des événements $Z \rightarrow e^+e^-$. Ceci est l'objet de cette section. Les résultats sont obtenus avec un échantillon de données enregistré en 2012, correspondant à une luminosité intégrée égale à 5.25 fb^{-1} , et un échantillon de simulation d'événements Drell–Yan (DY), dont la description détaillée se trouve dans la Section 3.2, et sont extraits du rapport de stage de Pauline Soulet [38].

2.4.1 Efficacités par électron

La première étape est la mesure de l'efficacité de chaque branche du *trigger* pour un électron reconstruit et sélectionné hors ligne. Cette étude est réalisée en fonction du p_T des électrons, pour la région du *barrel* et celle des *endcaps*, pour les données et la simulation. Les événements $Z \rightarrow e^+e^-$ sont sélectionnés par le *trigger* T&P HLT_E1e20_CaloIdVT_CaloIsoVT_TrkIdT_TrkIsoVT_SC4_Mass50 et l'électron *tag* doit vérifier $p_T > 20 \text{ GeV}$ et des sélections sévères similaires à celles à la branche E1e20 du *trigger*.

Dans un premier temps, la sélection hors ligne appliquée à l'électron *probe* est similaire à celle du *trigger* diélectron, exceptée la sélection en p_T . Les deux électrons doivent être de signes opposés et former une masse invariante $60 < m_{e^+e^-} < 120 \text{ GeV}$. Une fois cette sélection effectuée et pour chaque branche du *trigger* diélectron, l'efficacité est mesurée par la fraction d'électrons *probe* associés géométriquement avec un candidat HLT sélectionné par cette branche ($\Delta R < 0.1$).

La Fig. 2.2 montre les efficacités ainsi mesurées en fonction du p_T des électrons pour les deux branches E1e17 et E1e8 du *trigger* diélectron, pour le *barrel* et les *endcaps*, avec les données et la simulation. Les efficacités de chaque branche ne sont pas complètement indépendantes entre elles ; en effet les candidats HLT sélectionnés par chaque branche ne sont présents uniquement si l'événement vérifie la sélection finale du *trigger*. Ainsi l'efficacité pour une branche comprend le fait qu'un autre candidat vérifie l'autre branche ainsi que la sélection en δz entre les deux candidats. La largeur de la transition en efficacité entre $\approx 0 \%$

et $\approx 100\%$, qui dans le cas idéal serait égale à 0 GeV , est significative de la résolution en impulsion transverse des électrons au HLT par rapport à la mesure hors ligne. La différence entre la mesure au HLT et hors ligne pour les observables d'identification et d'isolation joue également un rôle dans la forme de la courbe, en particulier après la transition et au plateau. La transition est légèrement plus large pour les *endcaps* que pour le *barrel*, et bien plus étendue pour la branche **E1e8** que **E1e17**. En effet, la différence entre la mesure d'impulsion transverse au HLT et hors ligne est la plus importante à bas p_T , où la complémentarité des mesures dans le trajectographe et le ECAL est exploitée hors ligne (et décrite en détails dans la Section 3.4), contrairement au HLT où seule la mesure dans le ECAL intervient. La simulation est en bon accord avec les données, avec toutefois une efficacité systématiquement légèrement meilleure que dans les données. La différence d'efficacité entre les données et la simulation au plateau est $< 1\%$ pour les deux branches.

Les efficacités mesurées pour des électrons *probe* passant des sélections similaires à celles du *trigger* apportent des éléments intéressants sur le comportement du *trigger*. Cependant, l'étude d'efficacité du *trigger* spécifique à l'analyse du Higgs en quatre leptons nécessite une sélection d'électrons *probe* identique à celle appliquée lors de l'analyse. Celle-ci, décrite en détails dans la Section 3.5, comprend une sélection d'identification issue d'une analyse multi-variée, d'isolation, et de réjection de photons de conversion. Elle représente un niveau de sélection bien plus relâché que celui du *trigger*.

La Fig. 2.3 montre les efficacités mesurées pour des électrons *probe* vérifiant les sélections de l'analyse du Higgs en quatre leptons (**H4l**) en fonction de leur p_T pour les deux branches **E1e17** et **E1e8** du *trigger* diélectron, pour le *barrel* et les *endcaps*, avec les données et la simulation. Comme attendu la transition est plus large que pour des électrons *probe* sélectionnés plus strictement, en particulier le plateau est atteint pour des p_T plus élevés et une inefficacité persiste au niveau du plateau. Cet effet provient intégralement de la sélection plus relâchée des électrons *probe* par rapport aux conditions du *trigger*, en particulier il a été vérifié qu'il ne provient pas de l'accroissement du bruit de fond. La simulation est en très bon accord avec les données pour la branche **E1e17** et légèrement inférieure aux données dans la zone de transition pour la branche **E1e8**. La différence d'efficacité entre les données et la simulation au plateau est $< 1.1\%$ pour la branche **E1e17** et $< 0.8\%$ pour la branche **E1e8**. Pour la sélection similaire au *trigger*, comme pour la sélection de l'analyse, les valeurs d'efficacité au plateau sont systématiquement meilleures pour la simulation que pour les données. Cette petite différence peut s'expliquer par les canaux dynamiquement masqués lors de la prise de données (car trop bruyants par exemple), dont la prise en compte n'est pas intégralement possible dans la simulation.

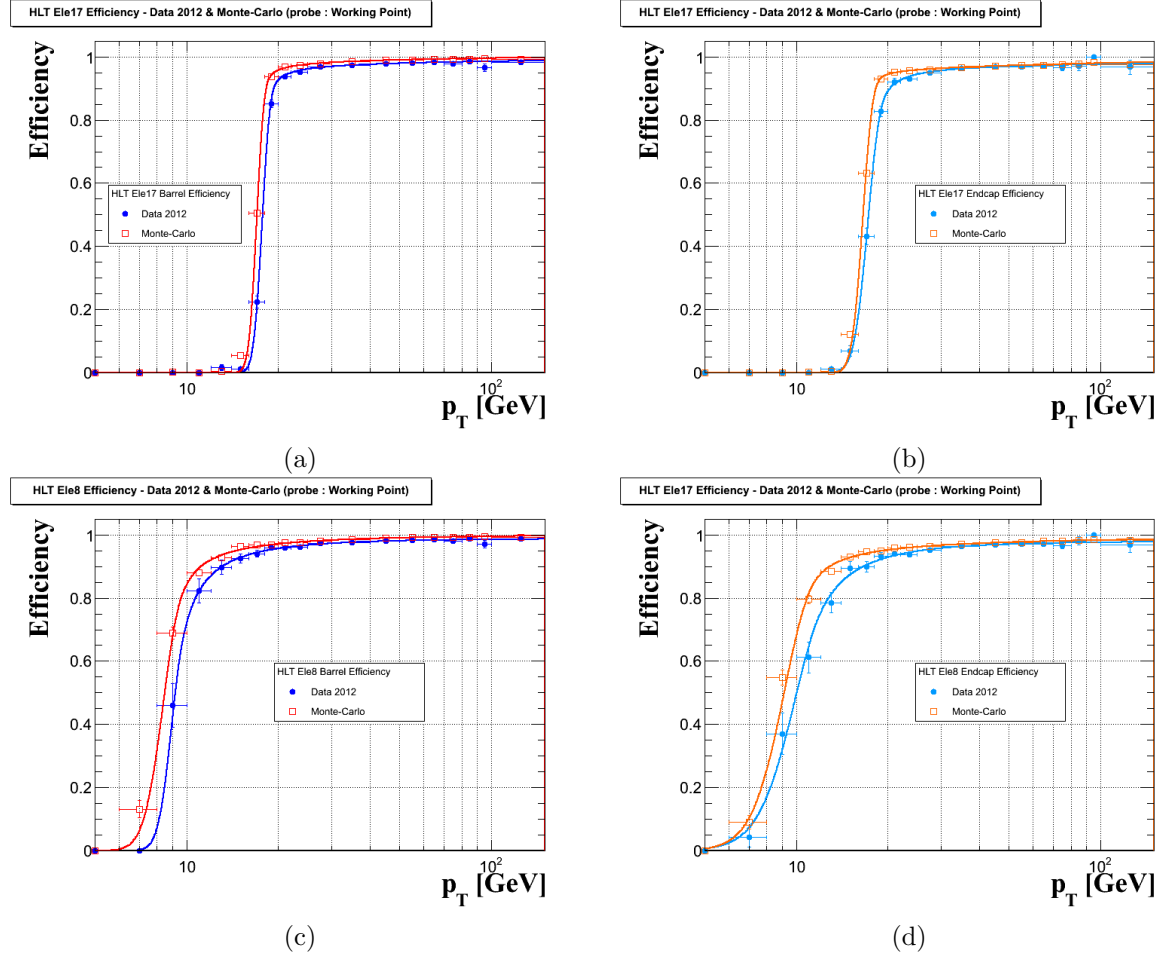


FIGURE 2.2 – Efficacités mesurées par la méthode T&P en fonction du p_T des électrons pour les branches (a) et (b) **Ele17**, (c) et (d) **Ele8** du *trigger* diélectron, pour (a) et (c) le *barrel* et (b) et (d) les *endcaps*, avec les données (en bleu) et la simulation (en rouge ou orange). La sélection appliquée sur les électrons *probe* correspond à la sélection du *trigger* diélectron.

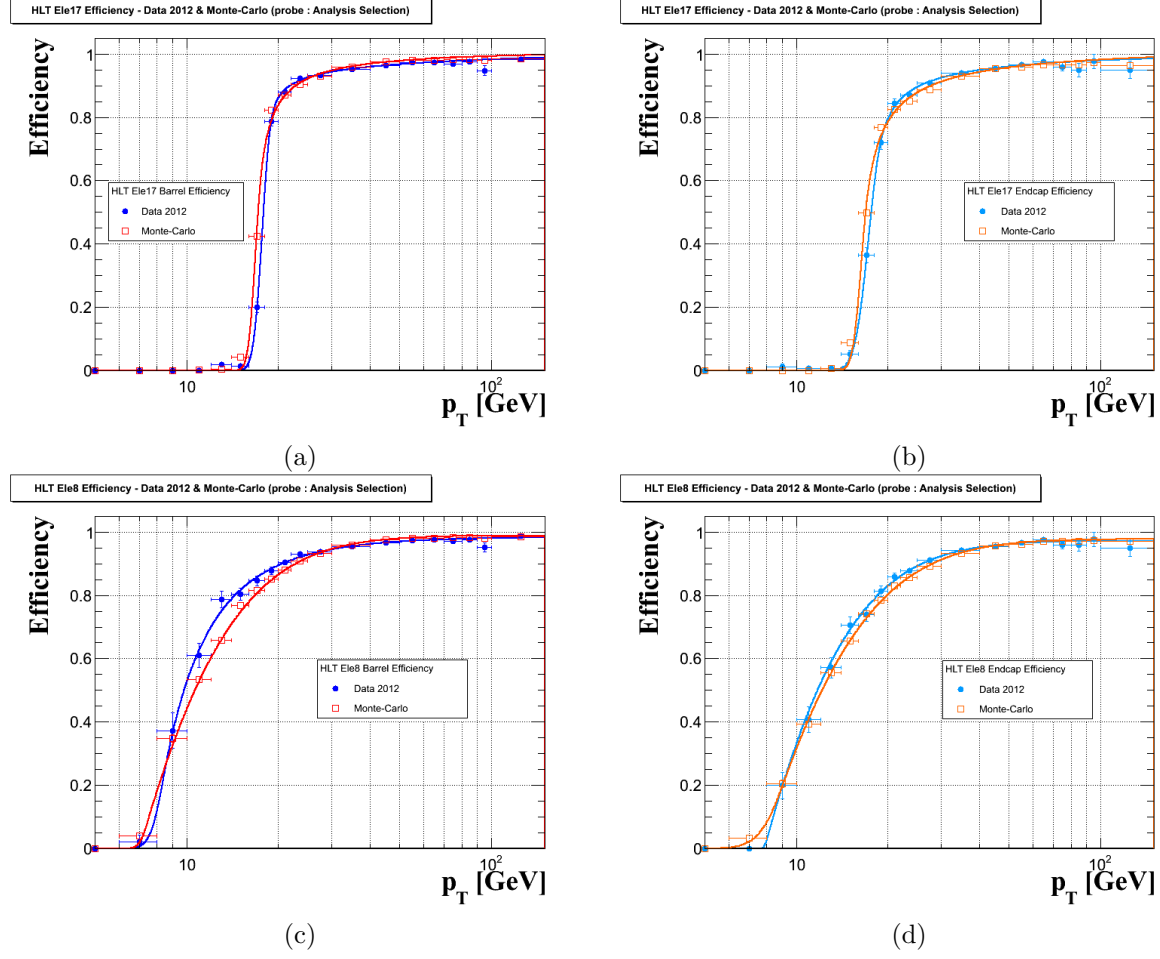


FIGURE 2.3 – Efficacités mesurées par la méthode T&P en fonction du p_T des électrons pour les branches (a) et (b) Ele17, (c) et (d) Ele8 du *trigger* diélectron, pour (a) et (c) le *barrel* et (b) et (d) les *endcaps*, avec les données (en bleu) et la simulation (en rouge ou orange). La sélection appliquée sur les électrons *probe* correspond à la sélection de l'analyse du boson de Higgs en quatre leptons.

Les informations sur les courbes d'efficacités sont données dans la Table 2.6 pour les deux sélections d'électrons *probe* (*trigger* et $H4\ell$) : les valeurs de p_T pour les efficacités à 50 et 95 %, et la valeur de l'efficacité à $p_T = 100$ GeV.

Branche, sélection	Région	Type de données	$p_T@50\%$ (GeV)	$p_T@95\%$ (GeV)	Efficacité @100 GeV (%)
Ele17, <i>trigger</i>	<i>barrel</i>	données	17.8	22.1	98.5
		MC	17.0	19.4	99.5
	<i>endcaps</i>	données	17.3	27.0	97.6
		MC	16.6	21.4	97.9
Ele8, <i>trigger</i>	<i>barrel</i>	données	9.2	18.4	98.5
		MC	8.5	14.7	99.5
	<i>endcaps</i>	données	10.0	26.0	97.6
		MC	9.1	19.5	98.3
Ele17, $H4\ell$	<i>barrel</i>	données	17.9	33.5	98.2
		MC	17.2	30.8	99.3
	<i>endcaps</i>	données	17.7	40.9	97.9
		MC	16.9	44.6	98.2
Ele8, $H4\ell$	<i>barrel</i>	données	9.8	32.5	98.0
		MC	10.7	31.2	98.8
	<i>endcaps</i>	données	11.8	39.4	97.1
		MC	12.2	41.3	97.6

TABLE 2.6 – Caractéristiques des courbes d'efficacité des deux branches du *trigger* diélectron : valeurs de p_T pour les efficacités à 50 et 95 %, et valeur de l'efficacité à $p_T = 100$ GeV.

2.4.2 Efficacités du *trigger* diélectron pour les événements Higgs en quatre électrons

Une fois les efficacités par lepton et par branche mesurées, en déduire l'efficacité du *trigger* pour un événement Higgs constitue une étape délicate. Pour cela, des événements de boson de Higgs simulés avec une masse $115 < m_H < 600$ GeV sont utilisés et sélectionnés avec quatre électrons vérifiant la sélection de l'analyse (décrite en détails dans la Section 4). Puis, à l'aide des courbes d'efficacité par électron vérifiant la sélection $H4\ell$, deux efficacités (une pour chaque branche) sont attribuées à chaque électron en fonction de son η et son p_T . L'électron est considéré vérifier la sélection HLT d'une branche si l'efficacité correspondante est supérieure au résultat d'un tirage aléatoire uniforme entre 0 et 1. Enfin, l'événement est considéré vérifier le *trigger* diélectron si les deux branches sont chacune vérifiées, et ce par deux électrons distincts. Cette estimation est effectuée à l'aide des courbes d'efficacité mesurées pour les données et pour la simulation. Les efficacités ainsi obtenues sont montrées

dans la Fig. 2.4 en fonction de m_H (courbes rouge et bleue). La différence entre les efficacités obtenues avec les courbes par électron pour la simulation d'une part et pour les données d'autre part montrent une différence négligeable, $< 0.1\%$ pour toute la gamme en masse. En effet, les différences entre les données et la simulation observées précédemment pour les courbes d'efficacité par électron sont très largement réduites par la présence de quatre électrons dans l'événement pouvant participer au *trigger*.

Cependant, ces efficacités comparées à l'estimation directe par la simulation (Fig. 2.4, courbe verte) montrent une nette différence, d'autant plus importante que la masse du boson de Higgs est basse, valant 3.5% pour $m_H = 125\text{ GeV}$. Cette différence peut être expliquée par trois effets principaux. Le premier provient de l'utilisation d'événements de bosons Z pour estimer les courbes d'efficacité par électron. Or, ceux-ci comprenant deux électrons ont une topologie dans le détecteur bien différente de celle des événements de Higgs en quatre leptons, en particulier pour les leptons proches pouvant entrer dans les cônes d'isolation les uns des autres. Le deuxième est le biais introduit par la méthode de T&P et en particulier par les sélections appliquées, qui améliorent artificiellement l'efficacité de par leur sévérité. Enfin, un dernier effet provient de la méthode simple d'estimation de l'efficacité d'un événement Higgs à partir des efficacités individuelles, celle-ci utilisant l'hypothèse non entièrement correcte de l'indépendance des électrons entre eux vis-à-vis des branches du *trigger* diélectron. Ces trois effets sont illustrés dans la même Fig. 2.4, où la courbe orange clair montre les efficacités évaluées à partir des courbes d'efficacité par électron obtenues avec la méthode de T&P appliquée à des événements de Higgs simulés, et la courbe orange foncé montre pour ces mêmes événements les efficacités évaluées à partir des courbes d'efficacité par électron, sans application de la méthode de T&P.

Concernant l'évaluation de l'erreur systématique liée à la différence de l'efficacité du *trigger* entre les données et la simulation pour l'analyse du boson de Higgs en quatre leptons, une approche pourrait considérer qu'il n'y a pas de raison pour que la différence négligeable observée avec la mesure précédemment décrite (courbes bleue et rouge) ne reste pas tout aussi négligeable pour l'efficacité réelle (courbe verte) et ainsi assigner une erreur systématique nulle. Cependant l'approche adoptée est plus conservatrice, assignant une erreur systématique de 1.5% couvrant la différence au plateau observée pour les courbes individuelles d'efficacité [27].

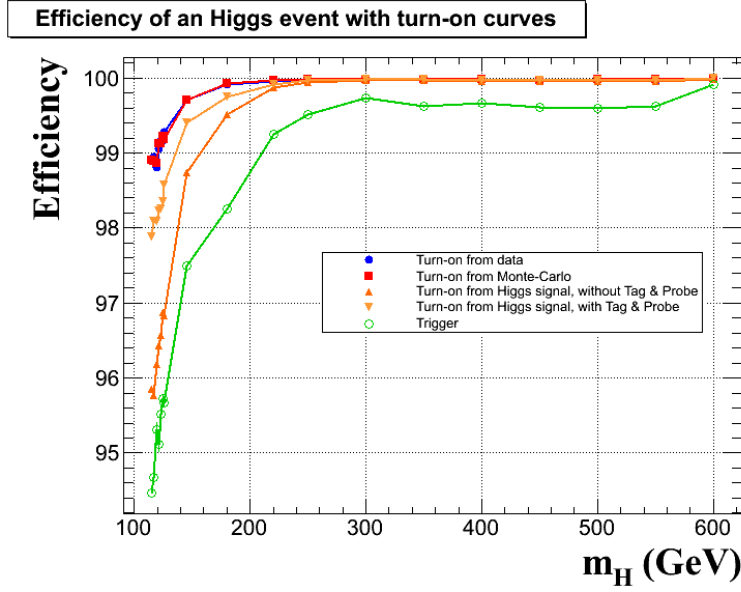


FIGURE 2.4 – Efficacités du *trigger* diélectron pour des événements de bosons de Higgs en quatre électrons simulés, en fonction de m_H , à partir des courbes d’efficacité individuelles par électron. Ces dernières sont estimées soit par T&P sur des événements Z pour les données (courbe bleue) et la simulation (courbe rouge), soit sur les événements Higgs simulés par T&P (courbe orange clair) ou sans T&P (courbe orange foncé). Les efficacités sont comparées avec la simulation directe du *trigger* sur les événements Higgs simulés (courbe verte).

Ce deuxième chapitre, dédié au déclenchement pour les événements de bosons de Higgs en quatre leptons, a décrit une étape déterminante pour le sujet de ce mémoire. À ce stade, nous avons donc enregistré les bosons de Higgs se désintégrant en quatre leptons, ainsi que d'autres événements, tels que des bosons Z se désintégrant en leptons, indispensables pour la suite de l'analyse. Le prochain chapitre est consacré à une étape fondamentale, et qui constitue le corps principal de ce mémoire, la reconstruction et l'identification des électrons.

Chapitre 3

Reconstruire

Sommaire

3.1	Défis	44
3.2	Échantillons de données et de simulation	45
3.3	Reconstruction algorithmique des électrons	47
3.3.1	Agrégation d'énergie dans le ECAL	47
3.3.2	Traces	49
3.3.2.1	<i>Seeding</i>	50
3.3.2.2	Reconstruction des traces	55
3.3.3	Agrégation PF de l'énergie des électrons	59
3.3.4	Association trace-cluster	60
3.3.5	Résolution des ambiguïtés	62
3.3.6	Estimation de la charge	62
3.4	Mesure de l'impulsion des électrons	66
3.4.1	Classification	66
3.4.2	Mesure d'énergie dans le ECAL	67
3.4.2.1	Énergies individuelles dans les cristaux	67
3.4.2.2	Corrections de l'énergie du SC	68
3.4.2.3	Échelle et résolution en énergie du SC	71
3.4.3	Combinaison de l'énergie du SC et de l'impulsion de la trace	72
3.4.4	Échelle et résolution en impulsion finales	73
3.5	Sélection des électrons	78
3.5.1	Identification	78
3.5.2	Isolation	84
3.5.3	Réjection des photons convertis	87
3.5.4	Sélection des électrons pour l'analyse $H \rightarrow ZZ^* \rightarrow 4\ell$	88
3.6	Mesures d'efficacité et de probabilité de fausse identification	89
3.6.1	Efficacité de reconstruction	91
3.6.2	Efficacité de sélection	91
3.6.3	Probabilité de fausse identification	95

Nous avons dorénavant enregistré les événements qui nous intéressent parmi ceux produits dans les collisions et nous savons quelle signature nous recherchons. Il faut maintenant se donner les moyens d’obtenir cette signature, c’est-à-dire d’être capable de reconstruire, à partir des indices laissés dans les différents sous-détecteurs, ces fameux leptons et en particulier les électrons. C’est ce à quoi est dédié ce chapitre, ainsi qu’une bonne partie de mes onze années sur CMS. Lors de la préparation du premier volume du “Physical Technical Design Report” (P-TDR) en 2005 [39], une refonte complète du logiciel de reconstruction des électrons a été effectuée afin d’améliorer sensiblement la reconstruction des électrons à bas p_T et la prise en compte du bremsstrahlung [40, 41]. Par la suite, les performances ont été continuellement améliorées et validées. Le présent chapitre est la traduction du papier électrons “legacy” de CMS [42] pour lequel j’ai été auteur et responsable, en binôme, agrémentée de quelques informations supplémentaires. Il traite d’aspects plus généraux que ceux spécifiques aux électrons dans l’analyse du Higgs en quatre leptons, mais ce canal a été principalement utilisé comme point de référence pour toutes les études inhérentes à la reconstruction et à la sélection des électrons.

3.1 Défis

Nous savons d’ores et déjà que notre signature est extrêmement exigeante. La rareté de notre signal conjuguée à la présence de leptons au nombre de quatre implique une reconstruction avec une efficacité maximale, et ce jusqu’à des impulsions transverses petites ($p_T < 5$ GeV), comme illustré par la Fig. 3.1 montrant la distribution en p_T des quatre leptons pour un boson de Higgs de masse 126 GeV. De plus, pour la sélection des événements mais également pour les mesures des caractéristiques du boson de Higgs, et en particulier de sa masse, une évaluation très précise de l’impulsion des leptons s’impose.

Par ailleurs, nous avons la chance de compter parmi nos événements des sources connues d’électrons, en particulier ceux issus de la désintégration du boson Z, comme illustré dans la Fig. 3.2 qui montre la distribution en masse invariante diélectron reconstruite dans les données enregistrées en 2012. Les événements contenant un boson Z sont considérablement utilisés pour accéder aux performances de la reconstruction et de la sélection des électrons.

Comme cela a été décrit dans le premier chapitre, deux détecteurs interviennent principalement dans la reconstruction des électrons : le trajectographe et le ECAL. La reconstruction des électrons dans CMS n’est pas une mince affaire. En effet, la conjugaison du champ magnétique intense et de la grande quantité de matière devant le ECAL implique que les électrons rayonnent une partie non négligeable de leur énergie par bremsstrahlung d’une part, et que d’autre part ces photons convertissent. Cela a pour conséquence un étalement et une structure complexes des électrons dans les détecteurs. D’un autre côté, la complémentarité du ECAL et du trajectographe apporte de nombreuses possibilités d’optimisation de la reconstruction et de la mesure des caractéristiques des électrons. La stratégie globale de re-

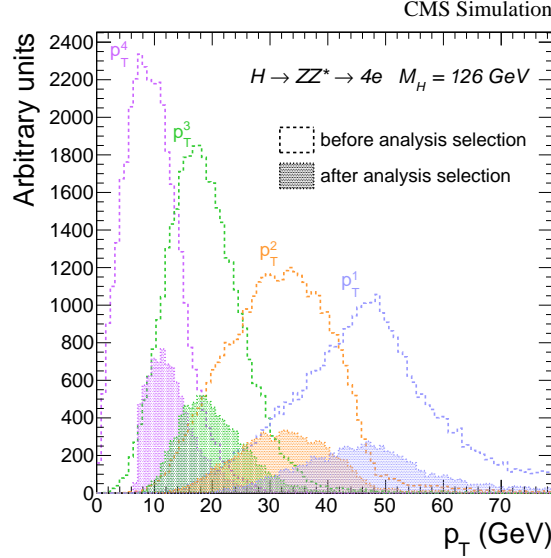


FIGURE 3.1 – Distributions en p_T des quatre leptons issus de la désintégration de bosons de Higgs (dans ce cas en quatre électrons) simulés avec une masse de 126 GeV, au niveau générateur à l’intérieur de l’acceptance en η du détecteur (histogrammes vides), et pour les événements sélectionnés à la fin de l’analyse du Higgs en quatre leptons, décrite dans le chapitre 4 (histogrammes remplis).

construction des électrons est l’association d’une trace dans le trajectographe et d’un amas d’énergie dans le ECAL (*cluster*). L’approche dite “standard” [40] est complétée par l’approche plus globale du “particle flow” (PF) [43,44].

3.2 Échantillons de données et de simulation

Les données utilisées pour les résultats exposés dans ce chapitre correspondent à une luminosité intégrée égale à 19.7 fb^{-1} [45], enregistrées par CMS à $\sqrt{s} = 8 \text{ TeV}$ durant l’année 2012. Ces données bénéficient de la calibration et des conditions d’alignement finales obtenues par les procédures décrites dans les Refs. [46,47]. Les performances de la reconstruction et de la sélection des électrons sont obtenues avec des événements sélectionnés par les *triggers* diélectron, décrits dans le chapitre 2, afin de collecter les électrons provenant principalement de désintégrations du Z, et dans une moindre mesure de résonances à basse masse. Les efficacités de reconstruction et de sélection sont étudiées grâce aux événements sélectionnés par les *triggers* spécifiques Tag&Probe, également décrits dans le chapitre 2, qui permettent la collection d’événements Z sans introduire de biais de sélection sur l’un des électrons. Enfin, les échantillons de bruit de fond faussement identifiés en électrons utilisent des événements

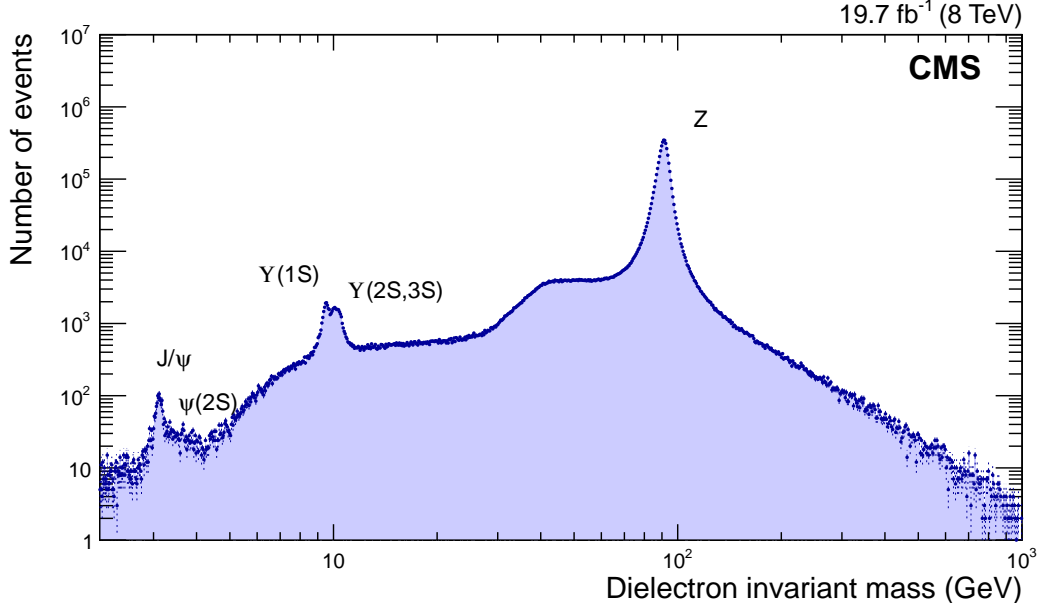


FIGURE 3.2 – Masse invariante diélectron reconstruite dans les données 2012.

$Z \rightarrow e^+e^-$ ou $Z \rightarrow \mu^+\mu^-$ contenant un jet additionnel et sont sélectionnés par les *triggers* diélectron ou dimuon.

Divers échantillons simulés sont exploités lors de l’optimisation des algorithmes de reconstruction et de sélection, de l’évaluation des efficacités ou des incertitudes systématiques. De façon générale les algorithmes de reconstruction sont optimisés sur des échantillons d’électrons simulés dos-à-dos avec des distributions uniformes en η et p_T , avec $1 < p_T < 100$ GeV. Concernant le signal, les comparaisons entre données et simulations ainsi que les études d’efficacités utilisent des événements Drell–Yan (DY) simulés, correspondant à une production quark + antiquark $\rightarrow Z/\gamma^* \rightarrow e^+e^-$. Ceux-ci sont générés à l’aide de MADGRAPH 5.1 [48], puis l’interface avec PYTHIA6.426 [49] est réalisée pour la description des radiations dans les états initial et final, de l’hadronisation des partons, et de l’événement sous-jacent (avec les paramètres PYTHIA tune Z2* [50]). Les événements additionnels dus à l’empilement d’interactions proton-proton lors d’un même croisement de faisceaux sont également générés et ajoutés. Ils sont au nombre de ≈ 21 en moyenne, correspondant à ≈ 15 vertex d’interactions reconstruits par événement considéré. Les signaux déposés par les événements dans le détecteur sont entièrement simulés en utilisant GEANT4 [51, 52] et reconstruits avec les mêmes algorithmes que pour les données. Une description réaliste des conditions du détecteur (alignement du trajectographe, calibration et alignement du ECAL, bruit électronique) est implémentée dans la simulation. En outre, lorsqu’une compréhension plus fine du détecteur est requise, une version de la simulation dépendante de la période de prise de données est utilisée pour suivre l’évolution de la réponse du détecteur en fonction

du temps. Cette version inclut l'évolution de la transparence des cristaux et du bruit dans le ECAL, et prend en compte par ailleurs les effets dus aux collisions dans une fenêtre en temps étendue par rapport au croisement de faisceaux de l'événement considéré.

3.3 Reconstruction algorithmique des électrons

Cette partie est consacrée à la reconstruction algorithmique des électrons et la mesure de certaines de leurs caractéristiques. Elle décrit en premier lieu la façon d'agréger les amas d'énergie dans le ECAL, puis la construction de la trace, leur association, et enfin la résolution des ambiguïtés, qui aboutit au candidat électron reconstruit final. La dernière section concerne l'estimation de la charge du candidat électron.

3.3.1 Agrégation d'énergie dans le ECAL

De façon générale, l'énergie déposée par un électron dans le ECAL est étalée sur quelques cristaux. Dans le cas de bremsstrahlung presque nul le long de la trajectoire de l'électron, cet étalement peut-être petit, par exemple des électrons d'un test en faisceaux d'une énergie de 120 GeV envoyés directement dans le centre d'un cristal déposent 97 % de leur énergie dans une matrice de 5×5 cristaux autour du point d'impact [25]. L'effet du bremsstrahlung peut être important : en moyenne 33 % de l'énergie d'un électron est rayonnée avant qu'il n'atteigne le ECAL pour une traversée minimale de matière ($|\eta| \approx 0$), et 86 % pour une traversée maximale ($|\eta| \approx 1.4$).

Afin de mesurer l'énergie initiale d'un électron (au moment de sa création), il est essentiel de collecter l'énergie de tous les photons émis le long de sa trajectoire, qui est principalement étalée le long de la direction ϕ par rapport à la direction initiale de l'électron. En général l'étalement en η est négligeable, à l'exception des électrons de très bas p_T ($p_T < 5$ GeV). Deux algorithmes de collection et d'agrégation d'énergie, appelés "Hybride" pour le *barrel* et "Multi5x5" pour les *endcaps*, sont utilisés à cet effet.

L'algorithme Hybride, illustré dans la Fig. 3.3 (a), exploite la géométrie du ECAL *barrel* et les propriétés de la forme de l'énergie déposée par l'électron dans le ECAL, en collectant l'énergie dans une petite fenêtre en η et une fenêtre étendue en ϕ [39]. Le point de départ est le cristal graine, défini comme celui contenant l'énergie maximale locale dans la région considérée et vérifiant $E_{Tseed} > E_{Tseed}^{\min}$. Autour de ce cristal graine, des dominos de cristaux de taille $\eta \times \phi = 5 \times 1$ sont considérés dans les deux directions en ϕ si leur énergie est supérieure à E_{domino}^{\min} , dans la limite de $\pm N_{steps}$ cristaux. Les dominos contigus sont alors groupés en *clusters*, chaque *cluster* distinct devant avoir un domino graine avec une énergie supérieure à $E_{seed-domino}^{\min}$ pour être collectés dans un *supercluster* (SC), agrégat correspondant à l'énergie de l'électron dans le ECAL. Les valeurs des seuils sont résumées dans la Table 3.1. Elles ont été initialement optimisées pour obtenir la meilleure résolution possible pour des électrons

avec $p_T \approx 15 \text{ GeV}$, quelques ajustements mineurs ont été apportés par la suite pour obtenir les bonnes performances actuelles sur une large gamme en p_T .

L'algorithme Multi5×5, illustré dans la Fig. 3.3 (b), est utilisé dans les *endcaps*, où les cristaux ne sont pas agencés dans une géométrie en $\eta \times \phi$. Le point de départ est constitué par des cristaux graines, d'énergie locale maximale par rapport à leurs quatre voisins directs et qui doivent vérifier $E_{T\text{seed}} > E_{TE\text{seed}}^{\min}$. Autour de ces cristaux graines, en commençant par celui de plus grande énergie transverse, l'énergie est collectée dans des *clusters* de 5×5 cristaux, qui peuvent en partie se recouper. Ces *clusters* sont alors groupés dans un SC si leur énergie transverse totale satisfait $E_{T\text{cluster}} > E_{T\text{cluster}}^{\min}$, dans une limite en η de $\pm\eta^{\text{range}}$ et une limite en ϕ de $\pm\phi^{\text{range}}$ autour de chaque cristal graine. Les valeurs de ces seuils sont également résumées dans la Table 3.1.

Dans les *endcaps*, les positions des *clusters* du SC, calculées par la moyenne des positions de leurs cristaux pondérée par les énergies correspondantes, sont extrapolées vers les plans du *preshower*. Le *cluster* d'énergie maximale est utilisé comme point de référence. La distance $\Delta\phi$ maximale entre les extrapolations des *clusters* et leur point de référence, étendue de 0.15 rad, est alors utilisée pour définir la fenêtre en ϕ de l'agrégation en énergie dans le *preshower*. Une fenêtre en η de taille $\Delta\eta = \pm 0.15$ par rapport au point de référence est également définie. Les énergies déposées dans le *preshower* à l'intérieur de ces fenêtres sont alors ajoutées à l'énergie du SC.

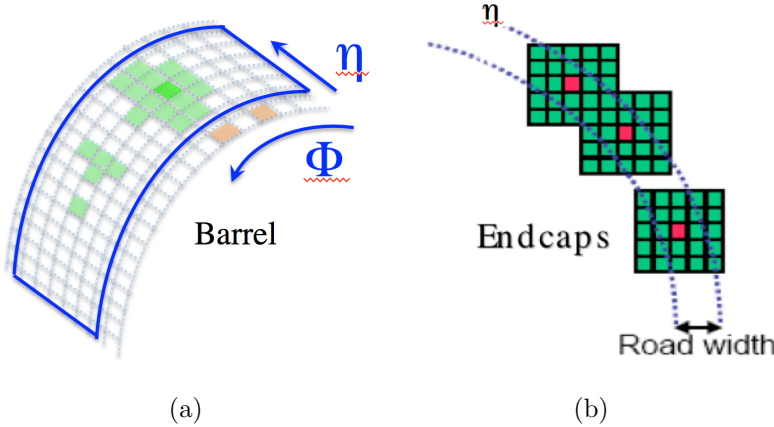


FIGURE 3.3 – Représentation schématique des algorithmes de collection et d'agrégation d'énergie, (a) pour l'algorithme Hybride dans le *barrel* et (b) pour l'algorithme Multi5×5 dans les *endcaps*.

L'énergie du SC correspond à la somme des énergies de tous les *clusters* constitutants. La position du SC est calculée par la moyenne des positions de ces *clusters*, pondérée par leurs énergies. La position de chaque *cluster* nécessite un traitement plus complexe. En effet, à

<i>Barrel</i>		<i>Endcaps</i>	
Paramètre	Valeur	Paramètre	Valeur
$E_{\text{Tseed}}^{\text{min}}$	1 GeV	$E_{\text{Tseed}}^{\text{min}}$	0.18 GeV
$E_{\text{domino}}^{\text{min}}$	0.1 GeV	$E_{\text{Tcluster}}^{\text{min}}$	1 GeV
$E_{\text{seed-domino}}^{\text{min}}$	0.35 GeV	η^{range}	0.07
N_{steps}	17 (≈ 0.3 rad)	ϕ^{range}	0.3 rad

TABLE 3.1 – Valeurs des différents seuils utilisés dans les algorithmes d’agrégation d’énergie dans le *barrel* (Hybride) et dans les *endcaps* (Multi5×5).

cause de la géométrie légèrement non-projective des cristaux et de la forme latérale de la gerbe électromagnétique, une simple moyenne des positions des cristaux pondérée par leurs énergies introduit un biais de la position de chaque *cluster* vers le centre de la gerbe, au lieu de la position d’entrée de l’électron dans le cristal. Une meilleure position est estimée d’une part en corrigeant la position de chaque cristal en fonction de l’estimation de la profondeur de la gerbe, et d’autre part en calculant la moyenne des positions des cristaux, pondérée par le logarithme de leurs énergies [39].

La Fig. 3.4 illustre, pour des événements $Z \rightarrow e^+e^-$ simulés, l’effet de l’algorithme d’agrégation en énergie sur l’efficacité de collection de l’énergie, en comparant l’énergie reconstruite en utilisant le SC avec celle reconstruite en utilisant une simple matrice de 5×5 cristaux autour du cristal le plus énergétique dans (a) le *barrel* et (b) les *endcaps*. Les queues de distributions à faibles valeurs du rapport de l’énergie reconstruite sur l’énergie générée E/E_{gen} sont significativement réduites en utilisant le SC.

Par ailleurs, dans le cadre de la reconstruction PF, un autre algorithme d’agrégation en énergie est introduit, dont le but est de reconstruire les gerbes électromagnétiques individuelles des particules. Les *clusters* PF sont reconstruits en agrégeant des cristaux contigus avec des énergies supérieures à deux déviations standards au-dessus du bruit électronique observé au début de la prise de données, autour d’un cristal graine d’énergie $E_{\text{seed}} > 230$ MeV dans le *barrel*, et $E_{\text{seed}} > 600$ MeV ou $E_{\text{Tseed}} > 150$ MeV dans les *endcaps*. Une différence notable par rapport à l’approche standard est la possibilité pour l’énergie d’un cristal d’appartenir à plusieurs *clusters*. Les *clusters* PF sont utilisés dans plusieurs étapes de la reconstruction des électrons.

3.3.2 Traces

La reconstruction de la trace des électrons a pour mission une collection efficace des points de mesures (*hits*) du trajectographe tout en préservant une détermination optimale des paramètres des traces, et ce pour une large gamme de fraction d’énergie émise par bremsstrahlung. Les traces des électrons peuvent être reconstruites grâce à la procédure standard

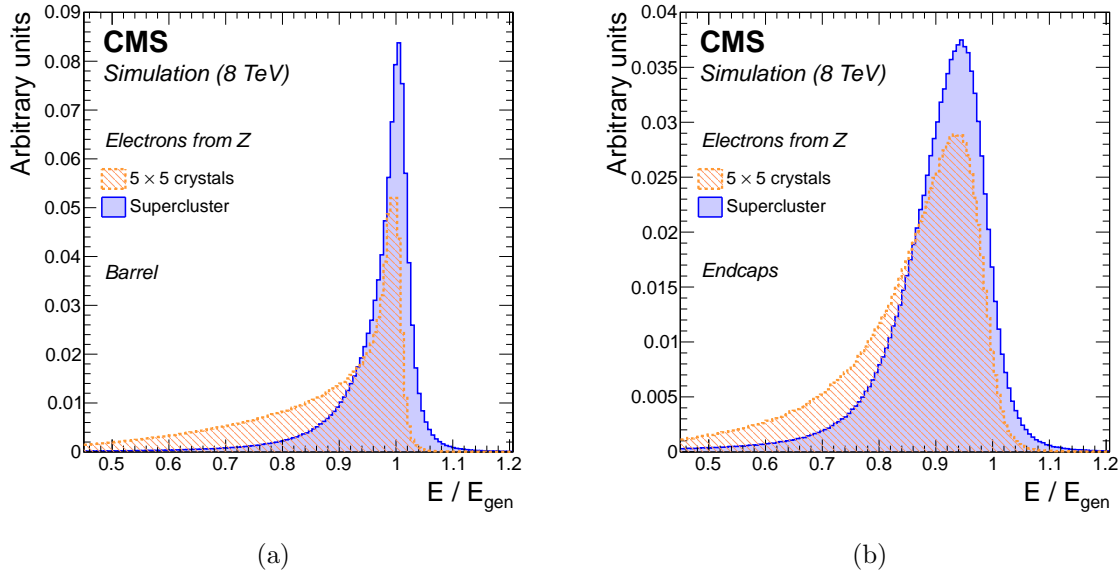


FIGURE 3.4 – Rapport de l’énergie reconstruite sur l’énergie générée des électrons issus d’événements Z simulés, en utilisant le SC comparé à la collection dans une matrice de 5×5 cristaux : (a) dans le *barrel*, (b) dans les *endcaps*. Aucune correction d’énergie n’est appliquée dans les distributions.

de filtre de Kalman (KF) utilisée pour toutes les particules chargées [23]. Cependant, cette procédure est inadaptée dans le cas des électrons à cause des pertes radiatives importantes dans la matière du trajectographe, et implique en général une efficacité réduite de collection des *hits* (perdus lors d’un changement de courbure important suite à du *bremsstrahlung*) ainsi qu’une mauvaise estimation des paramètres de la trace. Par conséquent, une procédure dédiée est utilisée pour la reconstruction des traces des électrons. Celle-ci étant très friande en temps, elle doit être initiée à partir de graines susceptibles de correspondre à des trajectoires initiales d’électrons, c’est l’objet du *seeding*.

3.3.2.1 *Seeding*

La première étape est donc la sélection des graines, soit les deux ou trois premiers *hits* dans le trajectographe à partir desquels la trace peut être initiée. Cette étape est d’importance primordiale car sa performance détermine l’efficacité de reconstruction des électrons. Deux algorithmes complémentaires sont utilisés et combinés. D’une part, le *seeding* “ECAL-driven” est initié par les SC dont l’énergie et la position sont utilisées pour estimer la trajectoire des électrons dans les premières couches de détection du trajectographe, et ainsi sélectionner les graines. D’autre part le *seeding* “Tracker-driven” utilise les traces générales, dont les extrapolations vers le ECAL sont associées aux *clusters* PF afin d’effectuer la sélection. Ces algorithmes, dont les paramètres ont été optimisés à l’aide de la simulation, ont été validés

avec succès avec les données enregistrées en 2010.

Les graines de traces sont formées par une association de paires ou de triplets de *hits* avec les vertex obtenus par les traces formées dans le trajectographe à pixels. Les premiers et deuxièmes *hits* des graines sont situés dans les couches du trajectographe à pixels dans le *barrel* (BPix) et les *endcaps* (FPix), ainsi que dans les disques externes à micro-pistes dans les *endcaps* (TEC) pour améliorer la couverture dans les régions à l'avant. Seule une sous-partie de ces graines donne finalement naissance à des traces.

ECAL-driven

L'énergie et la position de chaque SC sont utilisées pour extrapoler la trajectoire initiale des électrons au point de collision, en se basant sur le fait que la position moyenne des *clusters*, pondérée par leur énergie, correspond à la trajectoire hélicoïdale de l'électron initial propagée sans émission de radiation dans le champ magnétique de CMS jusqu'au ECAL. La propagation inversée de l'hélice à partir du SC est effectuée pour les deux hypothèses de charge de l'électron. L'intersection de l'hélice avec les couches les plus internes du trajectographe prédit les *hits* de la graine. Afin de limiter le nombre de graines faussement identifiées, une sélection est effectuée au niveau des SC : leur énergie transverse doit être suffisamment élevée, $E_T^{\text{SC}} > 4 \text{ GeV}$, et la quantité d'énergie contenue dans le HCAL à l'arrière suffisamment faible, $H/E_{\text{SC}} < 0.15$, E_{SC} étant l'énergie du SC et H la somme des énergies dans les tours du HCAL à l'intérieur d'un cône de taille $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} = 0.15$ autour de la direction du SC. Cette sélection permet un gain de temps de calcul important.

Une fois les *hits* de la graine prédits, la sélection est effectuée en comparant les positions prédites avec celles des graines en présence, et en choisissant celles qui sont suffisamment proches en ϕ et z (ou en distance radiale transverse r dans les régions à l'avant pour lesquelles les *hits* sont situés uniquement dans les disques). Les fenêtres de sélection pour le premier et le deuxième *hits* sont optimisées grâce à la simulation afin de maximiser l'efficacité de collection des graines tout en réduisant le nombre de graines faussement identifiées à un niveau gérable par les capacités de calcul, le facteur limitant étant le temps CPU total dédié à la reconstruction de la trace des électrons. L'efficacité globale du *seeding* ECAL-driven est $\approx 92\%$ pour des électrons issus de Z simulés.

Pour le premier *hit* les fenêtres de sélection sont larges et adaptées à l'incertitude de mesure de ϕ_{SC} et à la dispersion de la position du point de collision en z (σ_z), dépendante des conditions de faisceaux, et typiquement de l'ordre de 5 cm en 2012. De plus, la première fenêtre en ϕ dépend de E_T^{SC} afin de réduire le nombre de candidats faussement identifiés, et est asymétrique afin de prendre en compte l'incertitude sur l'énergie collectée dans les SC. Une fois le premier *hit* d'une graine sélectionné, il est utilisé pour affiner les paramètres de l'hélice et rechercher une compatibilité pour le deuxième *hit*, avec cette fois-ci des fenêtres plus réduites. Une graine est finalement sélectionnée si ses deux premiers *hits* sont associés

aux prédictions venant d'un SC.

Les valeurs des paramètres de la fenêtre de sélection pour le premier *hit* sont résumées dans la Table 3.2. Pour des électrons avec $5 < E_T^{\text{SC}} < 35 \text{ GeV}$, la taille de la fenêtre en ϕ ($\delta\phi$) est proportionnelle à $1/E_T^{\text{SC}}$. Le point donné à 10 GeV représente le point médian de la dépendance en E_T^{SC} . Les valeurs des paramètres de la fenêtre de sélection pour le deuxième *hit* sont résumées dans la Table 3.3.

$E_T^{\text{SC}} (\text{GeV})$	δz (BPix)	δr (FPix or TEC)	$\delta\phi$ (rad) (charge positive)
≤ 5	$\pm 5\sigma_z$	$\pm 5\sigma_z$	$[-0.075; 0.155]$
10	$\pm 5\sigma_z$	$\pm 5\sigma_z$	$[-0.046; 0.096]$
≥ 35	$\pm 5\sigma_z$	$\pm 5\sigma_z$	$[-0.026; 0.054]$

TABLE 3.2 – Valeurs des paramètres de la fenêtre de sélection pour le premier *hit*, en fonction de E_T^{SC} . σ_z est la déviation standard de la position en z de la collision, typiquement de l'ordre de 5 cm en 2012. Pour les candidats avec charge négative les mêmes tailles de fenêtre $\delta\phi$ sont utilisées mais avec des signes opposés.

δz (cm) (BPix)	δr (cm) (FPix)	δr (cm) (TEC)	$\delta\phi$ (rad) (BPix)	$\delta\phi$ (rad) (FPix or TEC)
± 0.09	± 0.15	± 0.2	± 0.004	± 0.006

TABLE 3.3 – Valeurs des paramètres de la fenêtre de sélection pour le deuxième *hit*.

La Fig. 3.5 montre la différence Δz_2 et $\Delta\phi_2$ entre les positions observées et prédites, (a) en z dans les pixels *barrel* BPix, et (b) en ϕ dans tous les sous-détecteurs du trajectographe, pour le deuxième *hit* de chaque graine de trace électron dans des événements $Z \rightarrow e^+e^-$ de données et simulés. Les distributions dans les données sont légèrement plus étalées que celles des simulations, l'effet étant plus important en $\Delta\phi_2$ et directement relié à la différence en résolution d'énergie entre les données et la simulation. Les fenêtres de sélection étant très larges, ce désaccord a une influence négligeable sur l'efficacité du *seeding*.

Tracker-driven

Le *seeding* “Tracker-driven”, partie intégrante de l'algorithme de reconstruction PF, apporte un complément en efficacité, en particulier pour les électrons de bas p_T ou non isolés, ainsi que pour des électrons dans la région de transition entre *barrel* et *endcaps*.

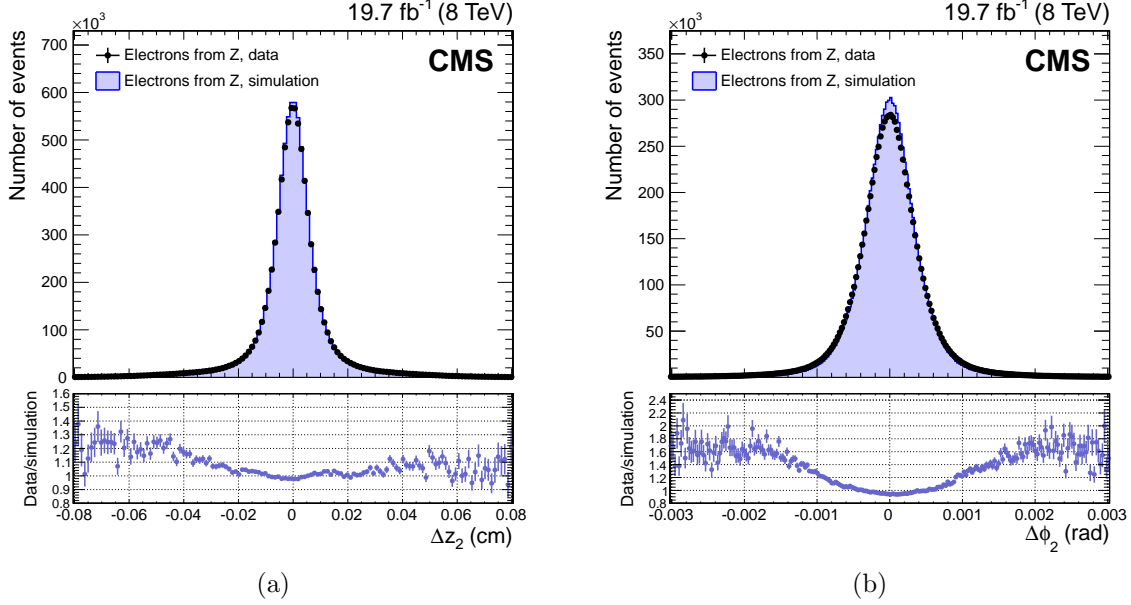


FIGURE 3.5 – Différences entre les positions observées et prédites des *hits* pour le deuxième *hit*, pour des électrons issus des désintégrations $Z \rightarrow e^+e^-$ dans les données et la simulation : (a) Δz_2 (BPix), (b) $\Delta \phi_2$ (tous les sous-détecteurs du trajectographe).

Le point de départ est constitué par les traces générales reconstruites en utilisant le KF. Dans le cas d'un électron dont le *bremsstrahlung* est négligeable, l'approche KF permet une reconstruction correcte et précise de sa trace. La trace KF bénéficie alors d'une collection de *hits* jusqu'au ECAL, d'une association avec le *cluster* PF le plus proche, et d'une mesure précise de son impulsion. En conséquence, la première étape de l'algorithme est la sélection des graines à l'origine des traces KF compatibles géométriquement et cinématiquement avec leur plus proche *cluster* PF. L'association géométrique est effectuée en comparant les directions en η et ϕ de la trace et du *cluster*, compte-tenu des incertitudes sur les positions du *cluster*. L'association cinématique compare l'énergie du *cluster* (E) et l'impulsion de la trace (p) et sélectionne la trace si $Ep_{\text{th}} < E/p < 3$, avec le seuil Ep_{th} égal à 0.75 pour les électrons avec $p_T > 6$ GeV, et relâché à 0.65 pour les électrons avec $2 < p_T < 6$ GeV pour lesquels la collection d'énergie dans le ECAL est plus difficile.

Pour les traces ne vérifiant pas les conditions ci-dessus, indiquant un possible *bremsstrahlung*, une seconde sélection est appliquée. En effet, la trace KF ne pouvant suivre le changement de courbure de la trajectoire de l'électron en cas de *bremsstrahlung*, soit elle arrête de collecter des *hits*, soit elle continue à les collecter et sa qualité est mauvaise, identifiée par une grande valeur du χ^2_{KF} . Les traces KF avec un petit nombre de *hits* ou un χ^2_{KF} important sont conséquemment réajustées en utilisant un filtre dédié aux électrons : le *Gaussian Sum Filter* (GSF) [53, 54], qui sera décrit dans la Section 3.3.2.2. Le nombre de *hits* et la qualité de la trace KF (χ^2_{KF}), la qualité de la trace GSF (χ^2_{GSF}), et les informations

concernant l'association géométrique et cinématique entre le *cluster* PF le plus proche et la trace sont alors exploités dans une analyse multivariée [55] pour déterminer la sélection de la graine de la trace comme une graine électron.

Seeding final

Les graines sélectionnées par les deux algorithmes sont combinées. L'efficacité finale de *seeding* prédite par la simulation est alors supérieure à 95 % pour des électrons issus des désintégrations $Z \rightarrow e^+e^-$. La Fig. 3.6 illustre cette efficacité de façon différentielle, en montrant également les efficacités individuelles pour les algorithmes ECAL-driven et Tracker-driven, en fonction de (a) p_T , et (b) η des électrons. L'efficacité augmente en fonction de p_T et les pertes d'efficacité en η proviennent principalement de la région de transition entre le *barrel* et les *endcaps*, et à la limite de l'acceptance du trajectographe. L'apport de l'algorithme Tracker-driven est bien visible à bas p_T et dans la région de transition entre le *barrel* et les *endcaps*. Par la suite les électrons dont la graine a été trouvée par l'algorithme ECAL-driven, respectivement Tracker-driven, sont simplement dénommés ECAL-driven, respectivement Tracker-driven. Une graine peut également être repérée par les deux algorithmes.

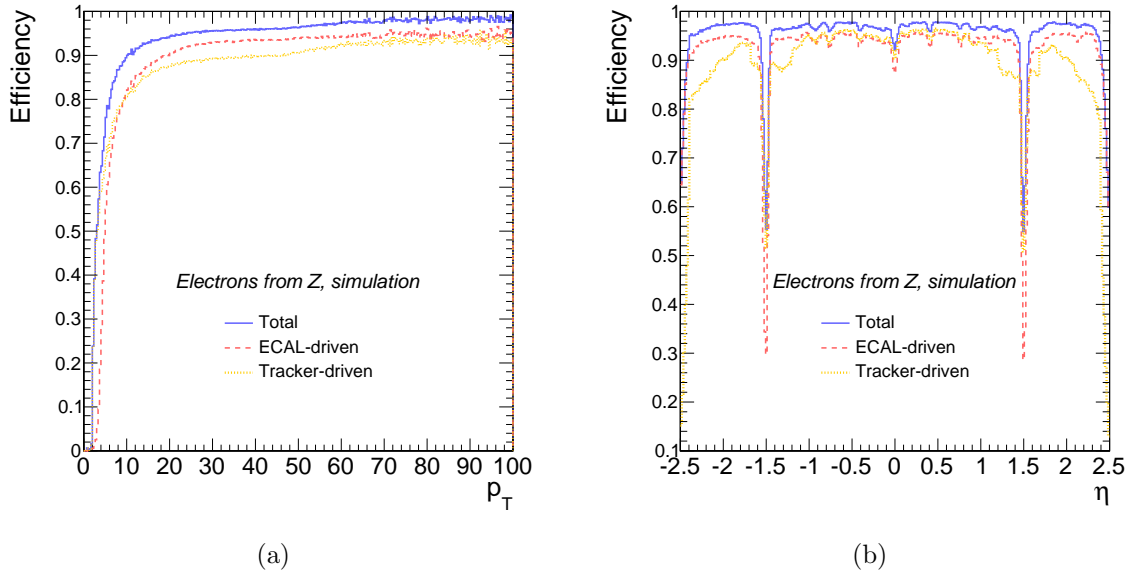


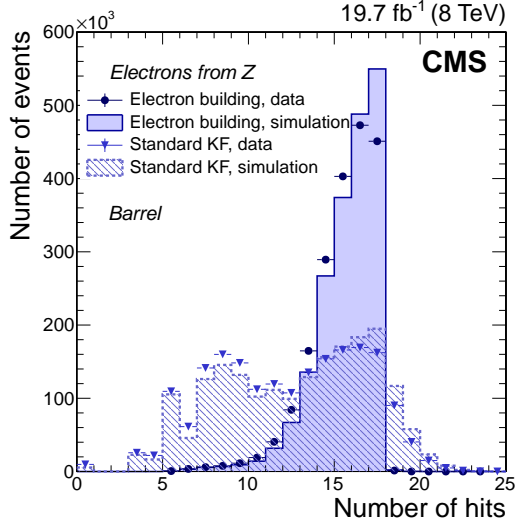
FIGURE 3.6 – Efficacité du *seeding* final (ligne continue), du *seeding* ECAL-driven (ligne en traitillé), et du *seeding* Tracker-driven (ligne en pointillé) pour les électrons issus d'événements $Z \rightarrow e^+e^-$ simulés : (a) en fonction de p_T , et (b) en fonction de η .

3.3.2.2 Reconstruction des traces

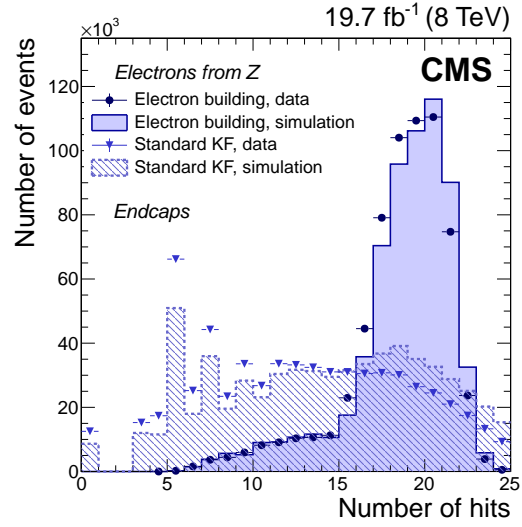
Les graines sélectionnées sont utilisées pour initier la construction des traces, suivie de leur ajustement. La construction de la trace est basée sur la méthode combinatoire KF, qui procède pour chaque graine de façon itérative à partir des paramètres de la trace mesurés à chaque couche, en incluant une par une les informations de chaque couche successive [23]. De cette façon la recherche de *hits* et l’ajustement de la trace sont réalisés simultanément. La perte d’énergie le long de la trajectoire est modélisée par une fonction de Bethe-Heitler, dédiée aux électrons. Afin de suivre la trajectoire de l’électron en cas de bremsstrahlung et dans le but de préserver une efficacité maximale, le critère de compatibilité (en terme de test de χ^2) entre les *hits* prédits et mesurés pour chaque couche est choisi très relâché. Lorsque plusieurs *hits* dans une couche sont compatibles avec les *hits* prédits, plusieurs trajectoires candidates sont créées et développées, avec une limite de 5 trajectoires candidates pour chaque couche. Un maximum d’un *hit* manquant est accepté par trajectoire candidate, et lorsque c’est le cas une pénalité de χ^2 lui est appliquée afin d’éviter d’inclure des *hits* provenant de photons de bremsstrahlung convertis dans la construction des traces électrons.

La Fig. 3.7 illustre, pour des électrons issus de Z dans les données et la simulation, le nombre de *hits* collectés grâce à cette procédure, comparé au nombre de *hits* collectés par la procédure KF standard utilisée pour toutes les autres particules chargées. La sélection d’événements $Z \rightarrow e^+e^-$ dans les données et la simulation requiert deux électrons isolés, non issus de conversion de photon, et avec $p_T > 20$ GeV. En outre, leur masse invariante doit satisfaire $|m_{e^+e^-} - m_Z| < 7.5$ GeV. Cette comparaison montre qu’en utilisant la procédure KF standard, les traces obtenues sont plus courtes qu’avec la construction dédiée. Le nombre de *hits* de la procédure KF standard est fixé à zéro lorsqu’aucune trace KF n’est associée avec la trace électron. La forme générale des distributions est bien reproduite par la simulation mais un désaccord est observé entre les données et la simulation, dû à une simulation imparfaite des senseurs actifs du trajectographe. L’effet éventuel de ce désaccord est pris en compte, en cas d’utilisation de cette observable pour l’identification, par des corrections apportées à l’efficacité extraites de l’analyse des événements $Z \rightarrow e^+e^-$.

Une fois tous les *hits* collectés, un ajustement GSF est effectué pour estimer les paramètres de la trace. À chaque couche, la distribution de probabilité d’énergie émise est approximée par une superposition de six fonctions gaussiennes, chacune propagée de la même façon vers la couche suivante. Un poids décrivant la probabilité associée est attribué à chaque composante (le nombre maximal de composantes à chaque couche est réduit à douze). À chaque point de mesure, deux estimations des paramètres de la trace sont exploitées en général : la moyenne pondérée de toutes les composantes (*mean*), ou la valeur la plus probable (*mode*). La première permet une estimation non biaisée en moyenne, et la seconde est piquée à la valeur générée et montre une largeur plus réduite pour l’essentiel de la distribution [40]. Ceci est illustré dans la Fig. 3.8, où le rapport p_T/p_T^{gen} entre les impulsions transverses mesurée et générée est comparé entre les deux estimations, pour des électrons issus de Z simulés. L’estimation *mode* est choisie pour tous les paramètres des traces des électrons.



(a)



(b)

FIGURE 3.7 – Nombres de *hits* collectés avec la procédure de construction de la trace dédiée aux électrons (points et histogrammes pleins), comparée à la procédure KF standard utilisée pour toutes les autres particules chargées (triangles et histogrammes hachurés), pour des électrons issus de Z dans les données (symboles) et la simulation (histogrammes) dans (a) le *barrel*, et (b) les *endcaps*.

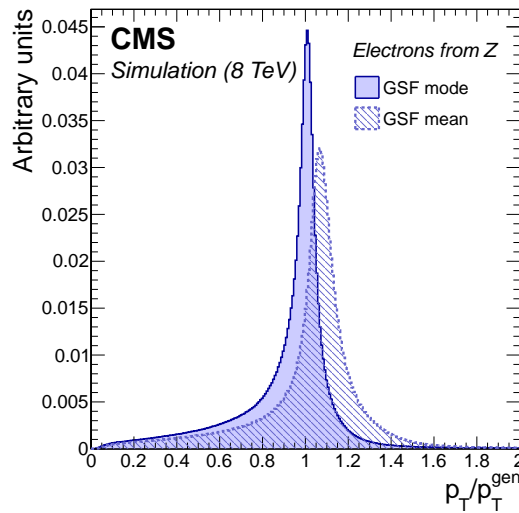


FIGURE 3.8 – Rapport p_T/p_T^{gen} estimé par la moyenne pondérée (*mean*, histogramme hachuré) comparé à celui estimé par la valeur la plus probable (*mode*, histogramme plein), pour des électrons issus de Z simulés.

Cette procédure de construction et d'ajustement de la trace des électrons permet de suivre la trajectoire des électrons jusqu'au ECAL, et ainsi d'estimer les paramètres de la trace à l'entrée du ECAL. La fraction d'énergie émise par bremsstrahlung est estimée grâce à l'impulsion de la trace mesurée à son point d'approche minimale de la collision (p_{in}), et à l'impulsion extrapolée à l'entrée du ECAL à partir des paramètres de la trace à la fin du trajectographe (p_{out}). Elle est définie par $f_{\text{brem}} = \frac{p_{\text{in}} - p_{\text{out}}}{p_{\text{in}}}$. Cette observable est utilisée dans la mesure de l'impulsion finale de l'électron et est également exploitée dans la procédure d'identification. Celle-ci est montrée dans la Fig. 3.9, pour des événements $Z \rightarrow e^+e^-$ de données et de simulation, ainsi que pour des candidats électrons faussement identifiés, provenant de bruit de fond de jets dans des données Z +jets, pour deux régions du *barrel* et deux des *end-caps*. Chaque distribution est normalisée à l'aire de la distribution correspondante dans les données $Z \rightarrow e^+e^-$. La sélection d'événements $Z \rightarrow e^+e^-$ dans les données et la simulation est identique à celle mentionnée précédemment : deux électrons avec $p_T > 20$ GeV, isolés et non issus de conversion de photon, et formant une masse invariante $|m_{e^+e^-} - m_Z| < 7.5$ GeV. Les électrons faussement identifiés sont obtenus en sélectionnant des électrons additionnels non isolés avec $p_T > 20$ GeV dans des événements de données contenant une paire de leptons identifiés avec une masse invariante compatible avec celle du Z , et une énergie transverse manquante < 25 GeV.

Lorsqu'un photon de bremsstrahlung est émis avant les trois premiers *hits* dans le trajectographe, impliquant une sous-estimation de p_{in} , ou lorsque le bremsstrahlung est très faible, avec pour conséquence des valeurs similaires pour p_{out} et p_{in} , l'observable f_{brem} peut atteindre des valeurs négatives. Dans la région la plus centrale du *barrel*, la quantité de matière traversée est relativement faible, et le maximum de la distribution de f_{brem} se situe à des valeurs faibles. En revanche, dans la région plus extérieure du *barrel* où la quantité de matière est plus importante, une grande part des électrons émettent des grandes fractions de leur énergie. Pour le bruit de fond, principalement composé de hadrons faussement identifiés en électrons, le maximum de la distribution de f_{brem} se situe à des valeurs très faibles. La contribution à grand f_{brem} d'électrons faussement identifiés, qui peut être observée dans les Figs. 3.9 (b), (c), et (d), est attribuée à des conversions de photon résiduelles intervenant très tôt, ainsi qu'à des interactions nucléaires dans le trajectographe.

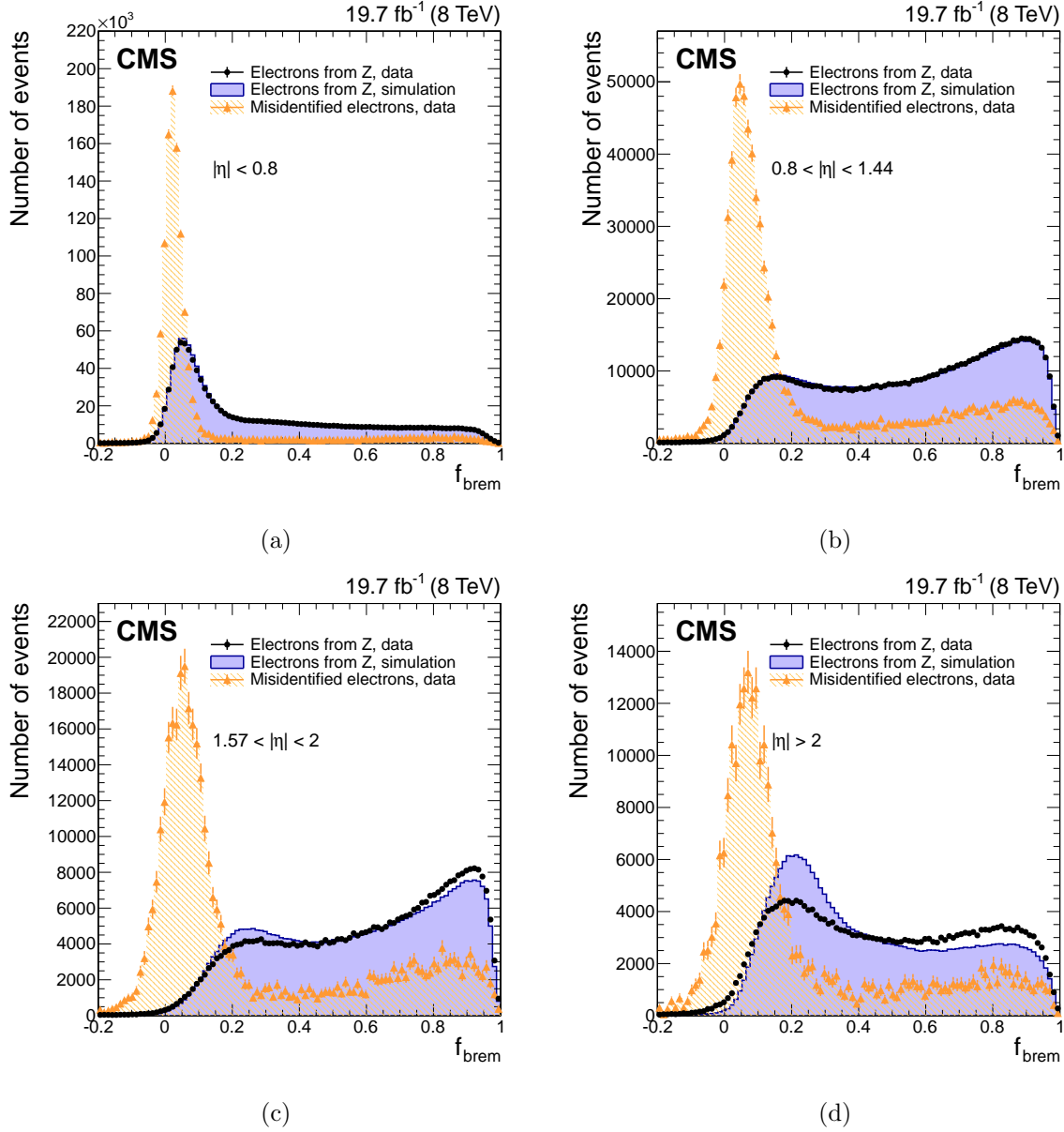


FIGURE 3.9 – Distributions de l’observable f_{brem} pour deux régions en $|\eta|$ du *barrel* et des *endcaps*, pour des électrons issus de Z dans les données (points) et la simulation (histogrammes pleins), et pour des électrons faussement identifiés dans les données (histogrammes hachurés). Chaque distribution est normalisée à l’aire de la distribution correspondante dans les données $Z \rightarrow e^+e^-$. (a) *barrel* central $|\eta| \leq 0.8$, (b) *barrel* externe $0.80 \leq |\eta| \leq 1.44$, (c) *endcaps* $1.57 \leq |\eta| \leq 2$, (d) *endcaps* $|\eta| \geq 2$.

Le désaccord observé entre les données et la simulation dans les *endcaps* est attribué à une description imparfaite du matériel dans la simulation. Il se trouve en effet que l'observable f_{brem} est un outil parfait pour accéder directement à la quantité de matériel dans le trajectographe. Ainsi, une comparaison directe de la valeur moyenne de f_{brem} entre les données et la simulation pour des tranches fines de η a permis l'identification d'une description imparfaite de matériel dans des régions spécifiques, en particulier du matériel manquant dans une région localisée à $|\eta| \approx 0.5$ et dans une région de services plus large à partir de $|\eta| \approx 0.8$ [56]. La différence observée entre les données et la simulation a permis de définir une version améliorée de la simulation de la géométrie du trajectographe pour les analyses futures. Cette étude est illustrée dans la Fig. 3.10, qui montre la comparaison de la mesure de la quantité de matériel avec des électrons issus de Z, par intervalle de η , entre les données et les simulations standard et améliorée. En ce qui concerne les analyses de 2011 et 2012, cette différence est prise en compte par des corrections apportées à l'échelle et la résolution en énergie, ainsi qu'aux efficacités de reconstruction et d'identification extraites de l'analyse des événements $Z \rightarrow e^+e^-$, comme discuté dans les Sections 3.4.4 et 3.6.

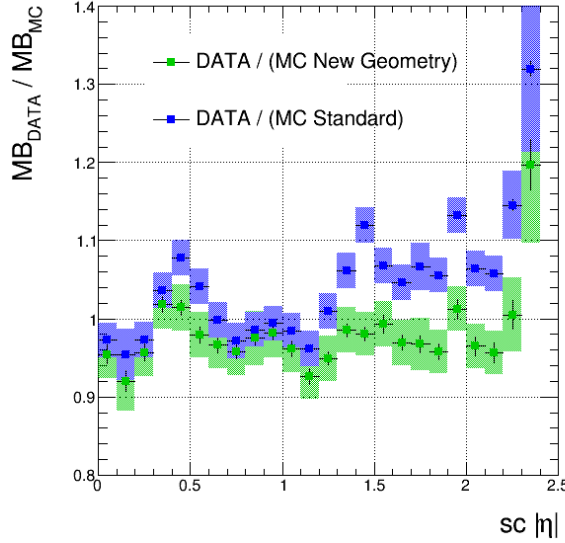


FIGURE 3.10 – Distributions du rapport de la quantité de matériel (*Material Budget*, MB) devant le ECAL, mesurée à l'aide de l'observable f_{brem} pour des électrons issus de Z, entre les données et la simulation standard (en bleu), et entre les données et la simulation améliorée (en vert) en fonction de la position en $|\eta|$ du SC des électrons [57].

3.3.3 Agrégation PF de l'énergie des électrons

L'agrégation PF des électrons est initiée par les traces GSF, indépendamment de l'algorithme à l'origine de leurs graines. Pour chaque trace GSF, plusieurs *clusters* sont assemblés, correspondant à l'électron à l'entrée du ECAL et aux photons de bremsstrahlung émis le

long de sa trajectoire. Le *cluster* PF correspondant à l'électron à l'entrée du ECAL est celui associé à la trace à la sortie du trajectographe. La plupart du matériel étant concentré au niveau des couches de détection du trajectographe, une droite tangente à la trajectoire est extrapolée depuis chaque couche jusqu'au ECAL, et chaque *cluster* PF associé à ces droites est ajouté au *cluster* électron PF. De cette façon la plupart des photons de bremsstrahlung sont collectés, à l'exception de certains photons convertis. Pour ceux-ci, une procédure spécifique sélectionne des traces KF secondaires grâce à un algorithme multivarié dédié et associe cinématiquement les *clusters* PF correspondants. En outre, pour les électrons isolés et ECAL-driven, tous les *clusters* PF associés géométriquement aux composants du SC Hybride ou Multi5×5 sont également ajoutés au *cluster* électron PF.

3.3.4 Association trace-cluster

Un candidat électron est construit à partir de l'association d'une trace GSF et de dépôts d'énergie (SC ou *cluster* électron PF) dans le ECAL. Pour les électrons ECAL-driven, le SC à l'origine de la graine est celui associé à la trace. Pour les électrons Tracker-driven, la trace est associée avec le *cluster* électron PF précédemment décrit.

Le critère de sélection de l'association trace-*cluster* est, comme la sélection de la graine, conçu dans le but de préserver une efficacité maximale, tout en réduisant la probabilité de fausse identification. Il est par conséquent assez relâché dans la direction la plus affectée par le bremsstrahlung. Dans le cas des électrons ECAL-driven, il s'agit des critères suivants d'association géométrique entre la trace GSF et le SC :

- $|\Delta\eta| = |\eta_{\text{SC}} - \eta_{\text{in}}^{\text{extrap}}| < 0.02$, η_{SC} étant la position en η du SC et $\eta_{\text{in}}^{\text{extrap}}$ la position en η de la trace extrapolée à partir de la position la plus interne jusqu'au point le plus proche du SC,
- $|\Delta\phi| = |\phi_{\text{SC}} - \phi_{\text{in}}^{\text{extrap}}| < 0.15$ rad, avec les définitions analogues en ϕ .

Pour les électrons Tracker-driven, une variable d'identification globale est définie à l'aide d'une technique multivariée. Elle combine les observables de la trace (cinématique, qualité, et comparaison avec la trace KF), les observables du *cluster* électron PF (forme et structure), et l'association entre les deux (géométrique et cinématique). Pour les électrons Tracker-driven uniquement, une sélection relâchée est appliquée sur cette variable d'identification. Pour les électrons à la fois Tracker-driven et ECAL-driven, la sélection est le résultat d'un OU logique entre les deux critères.

Pour des électrons issus de désintégrations de Z simulés, l'efficacité globale de reconstruction est $\approx 93\%$ et les efficacités différentielles sont montrées dans la Fig. 3.11 en fonction de p_T et η des électrons et du nombre de vertex d'interaction reconstruits. La mesure d'efficacité effectuée sur les données et comparée avec la simulation est montrée dans la Section 3.6.1. L'efficacité de reconstruction augmente rapidement en fonction de p_T pour atteindre 86% pour $p_T \approx 10$ GeV, et $\approx 95\%$ pour $p_T \approx 100$ GeV. Plusieurs réductions nettes d'efficacité apparaissent dans les régions en η correspondant aux zones peu ou pas instrumentées, entre

le *barrel* et les *endcaps*, entre les modules, et à la fin de l'acceptance du trajectographe à l'avant. L'efficacité en fonction du nombre de vertex d'interaction, par conséquent en fonction de l'empilement, est relativement stable avec une perte d'efficacité inférieure à 1 % pour une augmentation de 10 vertex d'interaction.

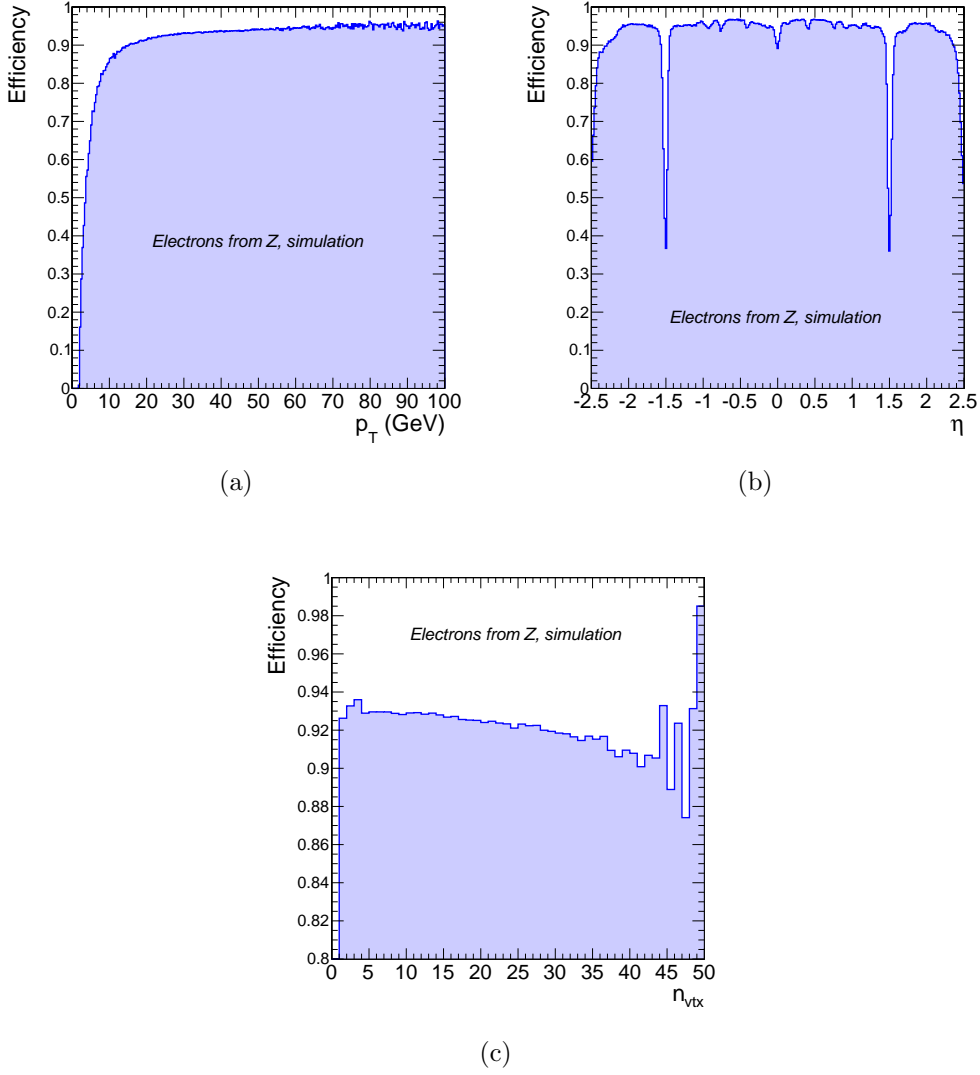


FIGURE 3.11 – Efficacité de reconstruction pour les électrons issus d'événements $Z \rightarrow e^+e^-$ simulés, en fonction de (a) p_T et (b) η des électrons, et (c) du nombre de vertex d'interaction reconstruits n_{vtx} .

3.3.5 Résolution des ambiguïtés

Dans le trajectographe, les photons de bremsstrahlung peuvent se convertir en paires e^+e^- qui donnent éventuellement lieu à des candidats électrons, en particulier dans la région $|\eta| > 2$, où les graines peuvent être situées dans les couches de détection des *endcaps* (TEC ou TID), loin du vertex d'interaction. Dans de telles topologies, d'un seul électron peuvent résulter de multiples traces. En particulier, lorsqu'un photon de bremsstrahlung emporte une importante fraction d'énergie, les *hits* correspondant à sa conversion peuvent être suffisamment proches de la position attendue pour la trace initiale de l'électron. De telles configurations impliquent la présence de candidats électrons ambigus, avec deux traces GSF proches associées au même *cluster* dans le ECAL (SC ou *cluster* électron PF, amalgamés par la suite dans le terme SC).

Afin de résoudre ces ambiguïtés et de choisir la trace GSF correspondant à l'électron considéré, les critères appliqués sont basés sur le fait qu'un photon a une probabilité faible de convertir juste après son point d'émission. Les intersections entre la trajectoire de la trace et les couches de détection actives les plus internes, comparées avec les *hits* présents dans la trace, permettent d'obtenir le nombre de *hits* internes manquants. Si deux traces GSF ambiguës ont un nombre différent de *hits* internes manquants, la trace choisie est celle présentant le plus petit nombre de *hits* internes manquants. Si le nombre de *hits* internes manquants est le même et que les deux traces candidates sont ECAL-driven, la trace choisie est celle avec le rapport E_{SC}/p le plus proche de la valeur unité, p étant l'impulsion de la trace au vertex d'interaction. Le même critère est appliqué dans le cas de deux traces candidates qui sont uniquement Tracker-driven. Si le nombre de *hits* internes manquants est le même et qu'une seule trace candidate est uniquement Tracker-driven, alors la trace ECAL-driven est choisie, à cause de la probabilité supérieure pour les traces Tracker-driven de correspondre à des conversions.

3.3.6 Estimation de la charge

La qualité de l'estimation de la charge est pénalisée par le bremsstrahlung suivi de la conversion des photons. En particulier, une conversion de photons de bremsstrahlung advenant tôt dans le détecteur donne lieu à des *patterns* très complexes et les produits de la conversion peuvent être indûment inclus dans l'ajustement des traces des électrons.

Un choix naturel pour l'estimation de la charge est donné par le signe de la courbure de la trace GSF. Cependant cette détermination peut être altérée par l'importante probabilité de mauvaise identification en présence de conversions, en particulier pour la région $|\eta| > 2$ où elle peut atteindre 10% pour des électrons issus de Z reconstruits sans sélection supplémentaire. Ceci est illustré dans la Fig. 3.13 montrant les probabilités de mauvaise identification de charge pour différentes méthodes en fonction de p_T et η des électrons reconstruits dans des événements Z simulés, sans sélection supplémentaire. Cette estimation est améliorée en la combinant avec deux autres mesures de la charge. La première est basée

sur la trace KF, associée à la trace GSF quand un *hit* est partagé dans la région la plus interne. La seconde est évaluée en utilisant la position du SC, et est définie par le signe de la différence en ϕ entre le vecteur joignant le point d'interaction à la position du SC et le vecteur joignant le point d'interaction et le premier *hit* de la trace GSF, comme illustré dans la Fig. 3.12.

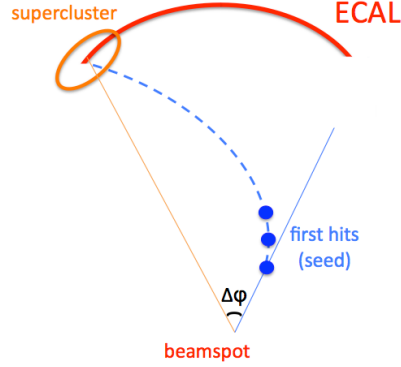


FIGURE 3.12 – Schéma de la différence $\Delta\phi$ entre le vecteur joignant le point d'interaction (*beamspot*) à la position du SC et le vecteur joignant le point d'interaction et le premier *hit* de la trace GSF [58].

La charge de l'électron est alors définie par la valeur donnée par deux estimations sur trois (méthode de la majorité). La probabilité de mauvaise identification de cet algorithme est prédite par la simulation à 1.5 % pour des électrons issus de Z reconstruits sans identification supplémentaire, offrant ainsi une amélioration globale de la probabilité de mauvaise identification d'un facteur ≈ 2 par rapport à la charge donnée par la courbure de la trace GSF uniquement. La probabilité de mauvaise identification est également réduite à très grand $|\eta|$, où elle est prédite inférieure à 7 % pour des électrons issus de Z. Une pureté accrue peut être obtenue en requérant que les trois estimations soient en accord (méthode de la sélection). La probabilité de mauvaise identification est alors inférieure à 0.2 % dans la partie centrale du *barrel*, inférieure à 0.5 % dans la partie externe du *barrel* et inférieure à 1.0 % dans les *endcaps*, ce qui peut être obtenu au prix d'une perte d'efficacité qui dépend de p_T , typiquement de l'ordre de 7 % pour les électrons issus de Z. Cet algorithme sélectif est utilisé principalement dans les analyses pour lesquelles l'estimation de la charge est cruciale, par exemple dans l'étude de l'asymétrie de charge dans la production inclusive de bosons W [59], où pour les recherche de supersymétrie utilisant des leptons de même charge [60]. La Fig. 3.13 montre les probabilités de mauvaise identification de charge pour chaque méthode en fonction de p_T et η des électrons reconstruits dans des événements Z simulés.

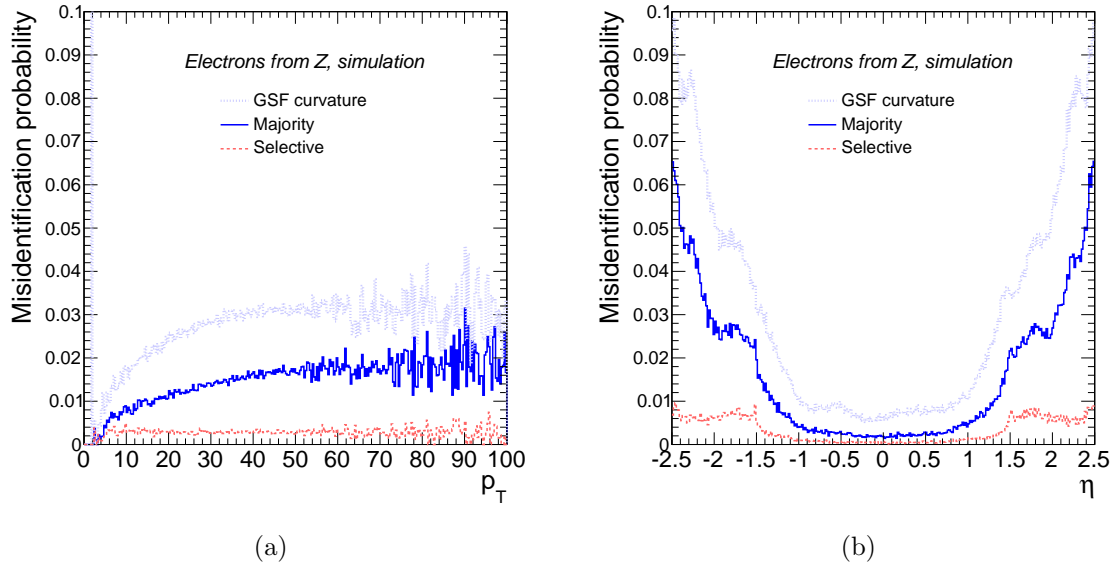


FIGURE 3.13 – Probabilité de mauvaise identification de charge en fonction de (a) p_T et (b) η des électrons reconstruits dans des événements Z simulés, sans sélection supplémentaire, pour la méthode de la courbure de la trace GSF (ligne en pointillé bleu clair), pour la méthode de la majorité (ligne continue bleu foncé) et pour la méthode de la sélection (ligne en traitillé orange).

La probabilité de mauvaise identification de la charge décroît fortement avec la sévérité des sélections d'identification, principalement grâce à la suppression des conversions de photons. La Fig. 3.14 montre la mesure de la probabilité de mauvaise identification de charge dans les données et la simulation, obtenue pour une sélection restrictive des électrons (correspondant au critère du HLT), pour les électrons issus de Z dans le *barrel* et dans les *endcaps*, et pour les méthodes de la majorité et de la sélection de l'estimation de la charge [61]. Ces valeurs sont obtenues en comparant le nombre de paires d'électrons de même charge et de charges opposées, issus d'un ajustement de la masse invariante diélectron. La probabilité de mauvaise identification est réduite drastiquement par rapport à celle évaluée au niveau de la reconstruction. Les rapports entre les données et la simulation (facteur d'échelle) sont également présentés sur la figure. Le désaccord entre les données et la simulation n'excède pas 20 % pour le *barrel* et les *endcaps*.

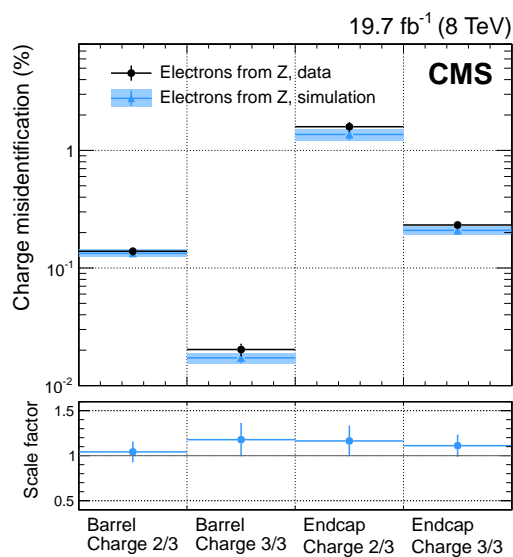


FIGURE 3.14 – Probabilité de mauvaise identification de charge pour une sélection restrictive d’électrons issus de bosons Z dans le *barrel* et dans les *endcaps*, dans les données (points) et la simulation (triangles), pour les méthodes de la majorité (2/3) et de la sélection (3/3) d’estimation de la charge. Les rapports entre les données et la simulation (facteur d’échelle) sont également présentés.

3.4 Mesure de l'impulsion des électrons

L'impulsion des électrons est estimée en combinant les mesures du trajectographe et du ECAL. Comme pour toutes les observables des électrons, sa mesure est particulièrement sensible au *pattern* des photons de bremsstrahlung et de leurs conversions. Afin d'obtenir la meilleure mesure possible, les électrons sont classifiés en fonction de leur schéma de bremsstrahlung, à l'aide d'observables sensibles à l'émission et la conversion des photons le long de la trajectoire de l'électron. L'énergie des SC est corrigée et calibrée, avant la combinaison entre les mesures du trajectographe et du ECAL.

3.4.1 Classification

Pour la grande majorité des électrons, la fraction de bremsstrahlung dans le trajectographe f_{brem} , définie dans le paragraphe 3.3.2.2, est complétée par la fraction de bremsstrahlung mesurée dans le ECAL, définie comme $f_{\text{brem}}^{\text{ECAL}} = \frac{E^{\text{PF}} - E_{\text{ele}}^{\text{PF}}}{E^{\text{PF}}}$, où E^{PF} et $E_{\text{ele}}^{\text{PF}}$ sont respectivement les énergies du *cluster* électron PF total et du *cluster* PF associé à l'électron à l'entrée du ECAL, correspondant aux énergies initiale et finale de l'électron. Le nombre de *clusters* total dans le SC est également exploité dans le processus de classification.

Les électrons sont classifiés dans les catégories suivantes :

- Les “*golden*” sont ceux avec peu de bremsstrahlung, fournissant par conséquent l'estimation la plus précise de l'impulsion. Ils sont définis par un SC avec un *cluster* unique et $f_{\text{brem}} < 0.5$.
- Les “*big-brem*” présentent un bremsstrahlung important émis en une seule étape, très tôt ou très tard le long de la trajectoire de l'électron. Ils sont définis par un SC avec un *cluster* unique et $f_{\text{brem}} > 0.5$.
- Les “*showering*” manifestent un bremsstrahlung important émis tout au long de la trajectoire de l'électron, et sont définis par un SC avec plusieurs *clusters*.

De plus, deux catégories spéciales sont définies. La première, nommée “*crack*”, correspond à des électrons dont le cristal graine du SC est adjacent à une limite en η du ECAL : entre les modules du *barrel*, entre le *barrel* et les *endcaps*, ou à l'extrême avant des *endcaps*. La seconde, appelée “*bad-track*”, est définie par les électrons dont la fraction de bremsstrahlung ECAL est significativement supérieure à la fraction de bremsstrahlung trajectographe ($f_{\text{brem}}^{\text{ECAL}} - f_{\text{brem}} > 0.15$), et permet d'identifier une trace de mauvaise qualité, en particulier dans la partie interne de la trajectoire.

La Fig. 3.15 (a) montre la fraction de la population des électrons dans les classes sus-nommées en fonction de $|\eta|$ (défini par rapport au centre de CMS), pour des électrons issus de désintégrations de bosons Z dans les données et la simulation. Les électrons *crack* ne sont pas montrés dans la figure mais complètent la proportion à l'unité. Les distributions pour les classes *golden* et *showering* reflètent la distribution en η du matériel devant le ECAL. Les données et la simulation sont en bon accord, excepté pour les régions en η connues pour

leur mauvaise modélisation du matériel, et pour $|\eta| > 2$, où le nombre de *clusters* des SC est surestimé par la simulation. Les proportions intégrées des électrons dans les différentes classes sont, respectivement pour les données et la simulation, 57.4 et 56.8 % pour les *showering*, 25.5 et 26.3 % pour les *golden*, 8.4 et 8.0 % pour les *big-brem*, 4.1 et 4.1 % pour les *bad-track*, et 4.6 et 4.7 % pour les *crack*. La Fig. 3.15 (b) montre les distributions du rapport de l'énergie du SC sur l'énergie générée (E_{gen}) pour les différentes classes. Les performances sont différentes en fonction des classes, et la qualité de l'estimation de l'énergie SC limitée pour les électrons avec un bremsstrahlung significatif. Cette estimation est améliorée grâce à des corrections additionnelles, discutées dans la section suivante.

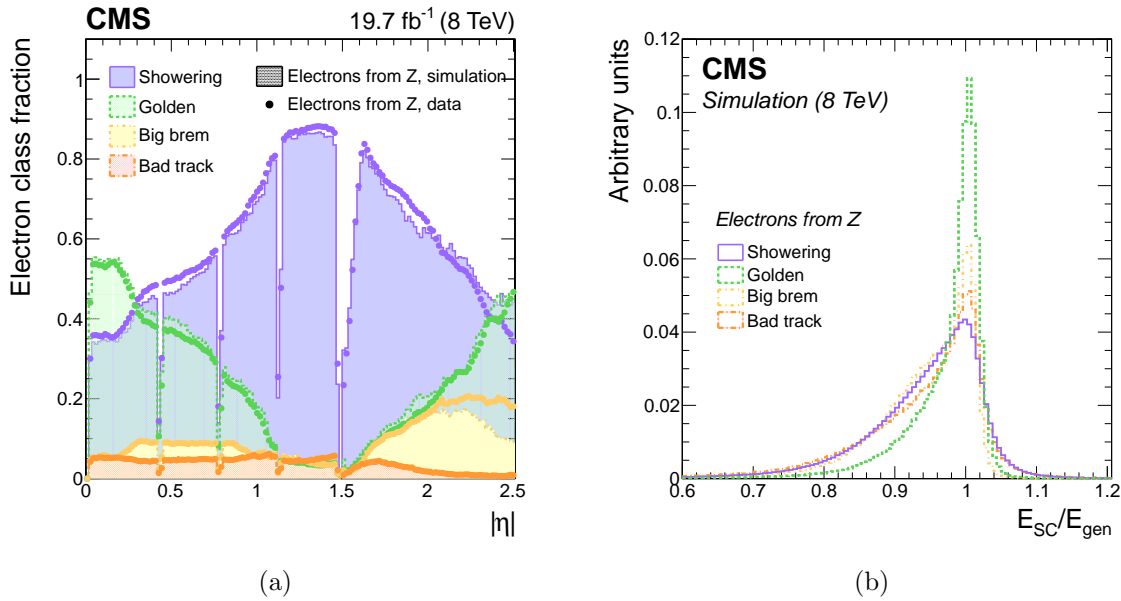


FIGURE 3.15 – (a) Fraction de la population des électrons issus de Z dans les différentes classes en fonction de $|\eta|$, pour les données (points) et la simulation (histogrammes). (b) Distribution de $E_{\text{SC}}/E_{\text{gen}}$ pour les différentes classes des électrons issus de Z simulés. Les électrons *crack* ne sont montrés dans aucune des figures.

3.4.2 Mesure d'énergie dans le ECAL

3.4.2.1 Énergies individuelles dans les cristaux

Avant l'étape de l'agrégation de l'énergie, la réponse en énergie individuelle de chacun des cristaux est calibrée à l'aide de plusieurs procédures [46]. L'amplitude de chaque cristal est reconstruite par une combinaison linéaire sur l'échantillonnage à 40 MHz du signal électrique issu du cristal. Cette amplitude est ensuite convertie en énergie en utilisant des facteurs mesurés séparément pour le *barrel*, les *endcaps*, et le *preshower*. Les effets sur la réponse du

cristal dus à l'irradiation sont corrigés grâce au système de contrôle laser du ECAL [62, 63], et les facteurs de correction sont validés dans les données avec la reconstruction de la masse invariante diélectron dans des événements $Z \rightarrow e^+e^-$ d'une part, et par le rapport de l'énergie du SC sur l'impulsion de la trace (E_{SC}/p) dans des événements $W \rightarrow e\nu$ d'autre part. Les coefficients d'inter-calibration entre les cristaux sont obtenus dans les données grâce à différentes méthodes : la symétrie en ϕ de la distribution d'énergie des événements de biais minimum à un η donné, la reconstruction des masses invariantes des événements $\pi^0 \rightarrow \gamma\gamma$, $\eta \rightarrow \gamma\gamma$, et $Z \rightarrow e^+e^-$, et enfin le rapport E_{SC}/p pour les électrons dans les événements $W \rightarrow e\nu$.

3.4.2.2 Corrections de l'énergie du SC

L'énergie du SC est obtenue en sommant les énergies individuelles des cristaux contenus dans le SC, avec dans le cas des *endcaps* celles correspondant au *preshower*. À ce stade, les effets principaux influençant l'estimation de l'énergie du SC sont liés à la collection d'énergie :

- fuites d'énergie en ϕ ou η en dehors du SC,
- fuites d'énergie dans les interstices entre les cristaux, modules, supermodules, et dans la région de transition entre le *barrel* et les *endcaps*,
- fuites d'énergie dans le HCAL à l'arrière du ECAL,
- pertes d'énergie lors d'interactions dans la matière avant le ECAL, et
- énergie additionnelle provenant des interactions d'empilement.

Les corrections nécessaires à l'énergie du SC sont obtenues grâce à une technique de régression multivariée [64, 65]. L'algorithme de régression est entraîné, séparément pour le *barrel* et les *endcaps*, à l'aide d'électrons simulés originellement avec un spectre uniforme en η et en p_T entre 5 et 300 GeV, puis repondérés afin d'accentuer la population à bas p_T . La cible de la régression est le rapport $E_{\text{gen}}/E_{\text{SC}}$. Les observables en entrée sont tout d'abord l'énergie du SC à corriger, et sa position en η et en ϕ , reliée à la quantité de matière intervenant devant le ECAL. Les observables de forme et de structure du SC et des *clusters* le composant (positions respectives, énergie et observables de forme) permettent d'accéder aux fuites en énergie en dehors du SC. Les fuites spécifiques aux interstices entre les modules, les supermodules et dans la région de transition *barrel-endcaps* sont explorées grâce à la position du cristal graine du SC. Afin de tenir compte des effets de fuite d'énergie entre les cristaux, la position du *cluster* graine relative à celle du cristal graine et les observables de forme de cascade électromagnétique sont utilisées. Le rapport H/E_{SC} (défini dans la Section 3.3.2.1) permet d'estimer la fuite d'énergie dans le HCAL. Les effets dus à l'empilement sont pris en compte grâce au nombre de vertex d'interaction reconstruits et la densité moyenne d'énergie ρ dans l'événement, définie comme la médiane de la distribution en densité d'énergie pour des particules à l'intérieur de chaque jet dans l'événement, reconstruit avec l'algorithme k_T [66, 67] avec un paramètre de distance égal à 0.6, $p_T^{\text{jet}} > 3 \text{ GeV}$ et $|\eta| < 2.5$.

La Fig. 3.16 montre la distribution du rapport entre l'énergie du SC corrigée par la régression et l'énergie générée, $E_{\text{SC}}^{\text{cor}}/E_{\text{gen}}$, pour deux catégories d'électrons simulés : électrons

de bas p_T ($7 \leq p_T < 10$ GeV) dans la partie centrale du *barrel*, et électrons de p_T modérés ($30 \leq p_T < 35$ GeV) dans la partie externe des *endcaps*.

Les distributions sont ajustées avec une fonction Crystal Ball [68] “double”. La fonction Crystal Ball est définie comme suit :

$$f_{CB}(x; \alpha, n, m_{CB}, \sigma_{CB}) = N \cdot \begin{cases} A \cdot \left(B - \frac{x - m_{CB}}{\sigma_{CB}} \right)^{-n}, & \text{pour } \frac{x - m_{CB}}{\sigma_{CB}} \leq -\alpha \\ \exp \left(-\frac{(x - m_{CB})^2}{2\sigma_{CB}^2} \right), & \text{pour } \frac{x - m_{CB}}{\sigma_{CB}} > -\alpha \end{cases} \quad (3.1)$$

où A et B sont des fonctions de α et n , et N un facteur de normalisation. Le but de cette fonction est de décrire à la fois le cœur gaussien (décrit par σ_{CB}) et les queues non gaussiennes (décrites par les paramètres n et α) de la distribution.

La fonction Crystal Ball double est une Crystal Ball modifiée avec les paramètres σ_{CB} , n , et α distincts pour des valeurs de x en dessous ou au-dessus de la position du maximum m_{CB} . La position du maximum et la déviation standard du cœur gaussien des distributions sont estimées respectivement grâce aux valeurs m_{CB} et σ_{CB} issues de l’ajustement. La déviation standard “effective” σ_{eff} , définie comme la moitié de l’intervalle minimum contenant 68.3 % des électrons autour de la position du maximum, est utilisée pour mesurer la résolution en tenant compte des queues non gaussiennes.

La position du maximum de la distribution $E_{SC}^{\text{cor}}/E_{\text{gen}}$ et la résolution effective de la distribution E_{SC}^{cor} sont montrées dans la Fig. 3.17, en fonction du nombre de vertex d’interaction reconstruits pour des électrons de p_T bas et modérés, dans le *barrel* et les *endcaps*. La position du maximum est affectée d’un biais inférieur à 1 %, indépendant du nombre d’interactions d’empilement, et reflet de la nature asymétrique de la distribution de la cible de la régression, E_{gen}/E_{SC} . La résolution effective est située dans l’intervalle 2 – 3 % pour les électrons de p_T modérés dans le *barrel*, et dans l’intervalle 7 – 9 % pour des électrons de bas p_T dans les *endcaps*. Elle se dégrade lentement avec le nombre d’interactions d’empilement.

Comparée à une paramétrisation standard pour la correction de E_{SC} en fonction de η , de la catégorie, et du E_T des électrons, la technique de régression multivariée apporte, en moyenne, une amélioration significative de ≈ 20 % sur la résolution, et jusqu’à ≈ 35 % dans les régions les plus à l’avant. De plus, le biais sur la position du maximum est réduit pour chaque classe et sur la gamme complète en η et p_T des électrons.

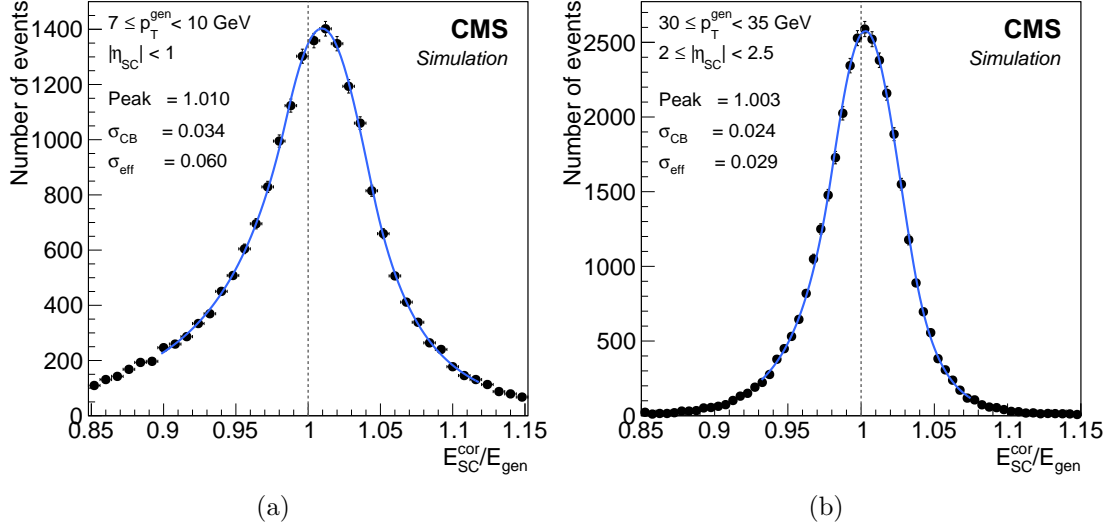


FIGURE 3.16 – Exemples de distributions du rapport entre l’énergie du SC, corrigée par régression, et l’énergie générée (E_{SC}^{cor}/E_{gen}) et leurs ajustements (par une fonction Crystal Ball double) pour deux régions en η et p_T : pour des électrons (a) avec $7 \leq p_T^{gen} < 10$ GeV et $|\eta_{SC}| < 1$, et (b) avec $30 \leq p_T^{gen} < 35$ GeV et $2 \leq |\eta_{SC}| < 2.5$, η_{SC} étant défini par rapport au centre de CMS. Les électrons sont générés avec des distributions uniformes en η et p_T .

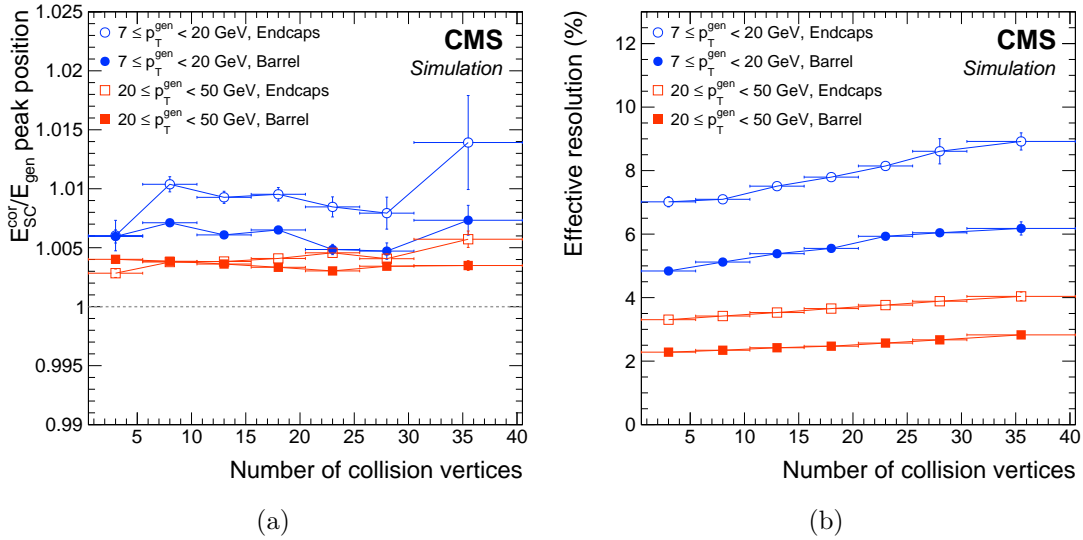


FIGURE 3.17 – (a) Position du maximum de E_{SC}^{cor}/E_{gen} , et (b) résolution effective de E_{SC}^{cor} , en fonction du nombre de vertex d’interaction reconstruits, pour des électrons dans le *barrel* (symboles pleins) et les *endcaps* (symboles vides) avec $7 \leq p_T^{gen} < 20$ GeV (cercles), et $20 \leq p_T^{gen} < 50$ GeV (carrés). Les électrons sont générés avec des distributions uniformes en η et p_T .

Un algorithme similaire de régression multivariée, basé sur les mêmes observables d'entrée, permet d'estimer les incertitudes sur l'énergie corrigée du SC, séparément pour les électrons du *barrel* et des *endcaps*, avec pour cible la différence absolue entre E_{gen} et l'énergie corrigée E_{SC} [65]. Un facteur multiplicatif égal à 1.253 est appliqué au résultat afin de transcrire la déviation absolue moyenne en une estimation, pour une distribution gaussienne, de la déviation standard. Comme attendu, les incertitudes relatives ainsi obtenues décroissent en fonction du p_T de l'électron, augmentent avec la quantité de matériel devant le ECAL et sont significativement plus conséquentes pour les électrons *showering* que pour les électrons *golden*.

3.4.2.3 Échelle et résolution en énergie du SC

Les corrections apportées à l'énergie du SC décrites précédemment sont basées sur la simulation. Les événements de données sont utilisés pour tenir compte des différences entre les données et la simulation pour les observables en entrée, et corriger les biais éventuels. Les corrections résiduelles sont petites ; d'une part, car l'énergie individuelle dans les cristaux est déjà calibrée, et d'autre part car la simulation des cascades électromagnétiques dans le ECAL est précise et inclut les incertitudes mesurées de l'inter-calibration entre les cristaux. La source principale de désaccord entre les données et la simulation est la description imparfaite du matériel du trajectographe dans la simulation, qui affecte différemment chaque catégorie d'électron. L'évolution de la transparence des cristaux et du bruit dans le ECAL durant la prise de données, lorsque non considérée par une simulation spécifique dépendante des périodes de prise de données, peut ajouter une différence entre les données et la simulation. Une autre source de désaccord pourrait être la sous-estimation des incertitudes de la calibration individuelle des cristaux. Enfin, une différence entre les géométries réelle et nominale du ECAL peut impliquer que les corrections discutées précédemment, obtenues avec des événements simulés avec la géométrie nominale, soient sous-optimales pour les données. Alors qu'il est maintenant compris qu'au moins un de ces effets contribue à la dégradation de la calibration, leurs ampleurs respectives ne sont pas intégralement connues. Plus de détails peuvent être trouvés dans la Ref. [56].

L'échelle d'énergie du SC est corrigée dans les données pour correspondre à celle de la simulation. Ces corrections sont obtenues grâce à des événements $Z \rightarrow e^+e^-$, en comparant la masse invariante diélectron entre les données et la simulation, pour quatre régions en $|\eta|$ et deux catégories d'électrons, sur 50 périodes de prise de données, en suivant la procédure décrite dans la Ref. [46]. Les régions en $|\eta|$ sont définies de la plus centrale à la plus à l'avant : *barrel* $|\eta| \leq 1$, *barrel* $|\eta| > 1$, *endcaps* $|\eta| \leq 2$, et *endcaps* $|\eta| > 2$. La variable R_9 , définie comme le rapport entre l'énergie reconstruite dans une matrice de 3×3 cristaux centrée sur le cristal le plus énergétique et l'énergie du SC, est utilisée pour accéder à la quantité de bremsstrahlung. La catégorie avec peu de bremsstrahlung est définie par $R_9 \geq 0.94$, et celle avec une quantité significative de bremsstrahlung par $R_9 < 0.94$. La masse invariante diélectron est reconstruite avec les énergies des SC et les directions mesurées par les traces. La distribution de masse invariante entre 60 et 120 GeV est ajustée à l'aide d'une fonction

Breit–Wigner convoluée avec une fonction Crystal Ball, pour les données et la simulation. Les corrections d’échelle, obtenues pour chaque catégorie par la différence entre les positions du maximum dans les données et la simulation, sont appliquées dans les données de telle sorte que les positions du maximum de la masse invariante soient en accord entre les données et la simulation, dans chaque catégorie. Ces corrections sont comprises entre 0.9880 et 1.0076 et leurs incertitudes entre 0.0002 et 0.0029.

L’estimation de la résolution en énergie du SC est également affectée par les mêmes sources de désaccord entre les données et la simulation. Cette fois-ci les corrections sont appliquées sur la simulation pour qu’elle corresponde aux données. Celles-ci sont indépendantes de la période de prise de données, et évaluées pour les mêmes catégories de $|\eta|$ et R_9 . L’énergie du SC est modifiée en appliquant un facteur issu d’une distribution gaussienne, centrée sur la valeur corrigée d’échelle et avec une déviation standard égale à $\delta\sigma^e$, correspondant à un terme constant additionnel dans la résolution en énergie. La valeur de $\delta\sigma^e$ est mesurée pour chaque catégorie d’électron à l’aide d’un maximum de vraisemblance entre les données et la simulation dont la résolution est modifiée. Ce terme constant additionnel de la résolution est compris entre $(0.92 \pm 0.03) \%$ dans la catégorie $|\eta| < 1$ et $R_9 \geq 0.94$, et $(2.90 \pm 0.03) \%$ dans la catégorie $|\eta| > 2$ et $R_9 < 0.94$. L’incertitude sur l’énergie du SC est augmentée conformément.

3.4.3 Combinaison de l’énergie du SC et de l’impulsion de la trace

La combinaison de l’énergie du SC (après toutes les corrections discutées précédemment) et de l’impulsion de la trace permet une amélioration substantielle de l’estimation de l’impulsion de l’électron (p_{comb}). En effet, pour des électrons d’énergie faible ou modérée ($\lesssim 15$ GeV), ou pour des électrons proches des interstices entre les régions actives de détection du ECAL, la mesure de l’impulsion de la trace est escomptée plus précise que celle de l’énergie du SC. Une technique de régression multivariée permet de définir le poids w qui multiplie l’impulsion de la trace dans la combinaison linéaire avec l’énergie du SC : $p_{\text{comb}} = w \times p + (1 - w) \times E_{\text{SC}}$. L’incertitude sur p_{comb} est calculée par la combinaison des incertitudes sur p et sur E_{SC} en utilisant le même poids.

La complémentarité des deux estimations dépend de la quantité de bremsstrahlung. L’énergie du SC corrigée et son incertitude relative, avec l’impulsion de la trace et son incertitude relative sont les observables d’entrée principales de la régression. L’adjonction du rapport E_{SC}/p avec son incertitude, et du rapport entre les deux incertitudes relatives de E_{SC} et p , introduit une information plus sophistiquée permettant l’optimisation de la performance de la régression. La classe de l’électron, et sa présence dans le *barrel* ou les *endcaps*, sont également incluses comme indicateurs de la qualité de l’électron et de la quantité de bremsstrahlung.

Après la combinaison, le biais dans la mesure de l’impulsion des électrons est réduit dans toutes les régions en η et pour toutes les classes d’électrons, à la faible exception des

électrons *showering* dans les *endcaps*, pour lesquels le biais devient très légèrement ($\approx 0.1\%$) pire. La Fig. 3.18 montre la résolution effective de l'impulsion de l'électron (en pourcents) en fonction de son p_T généré, après combinaison de E_{SC} et p , comparée à la résolution effective de l'énergie du SC corrigée, pour des électrons *golden* dans le *barrel* et *showering* dans les *endcaps*. L'amélioration est typiquement de $\approx 25\%$ pour des électrons avec $p_T \approx 15$ GeV dans le *barrel*, et atteint 50 % pour des électrons *golden* dans le *barrel* pour $p_T < 10$ GeV.

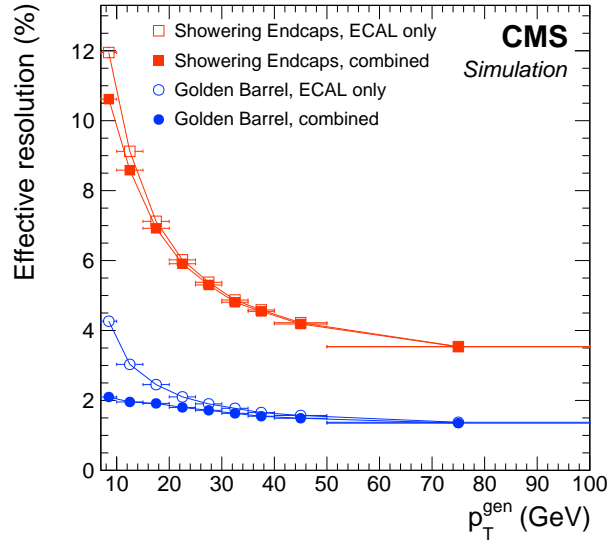


FIGURE 3.18 – Résolution effective de l'impulsion des électrons, après combinaison de E_{SC} et p (symboles pleins), comparée à celle de l'énergie du SC (symboles vides), en fonction du p_T généré de l'électron. Les électrons *golden* (cercles) et *showering* (carrés) sont montrés en exemple. Les électrons sont générés avec des distributions uniformes en η et p_T , et la résolution est montrée après correction.

L'amélioration de la résolution est significative pour tous les électrons dans le *barrel* à des énergies jusqu'à ≈ 35 GeV, comme illustré dans la Fig. 3.19 (a) qui montre la résolution effective de l'énergie du SC corrigée, de l'impulsion de la trace, et de leur combinaison, en fonction de l'énergie générée de l'électron. La Fig. 3.19 (b) représente la masse invariante reconstruite pour un boson de Higgs simulé avec une masse de 126 GeV dans le canal $H \rightarrow ZZ^* \rightarrow 4e$. La distribution de la masse reconstruite en utilisant l'énergie du SC corrigée est comparée à celle reconstruite après combinaison. La résolution effective est améliorée de 7 %, la résolution correspondant au cœur gaussien de 9 %.

3.4.4 Échelle et résolution en impulsion finales

L'échelle d'énergie obtenue dans les paragraphes précédents concerne l'énergie du SC dans des événements $Z \rightarrow e^+e^-$. Elle doit par conséquent subir des corrections additionnelles, en

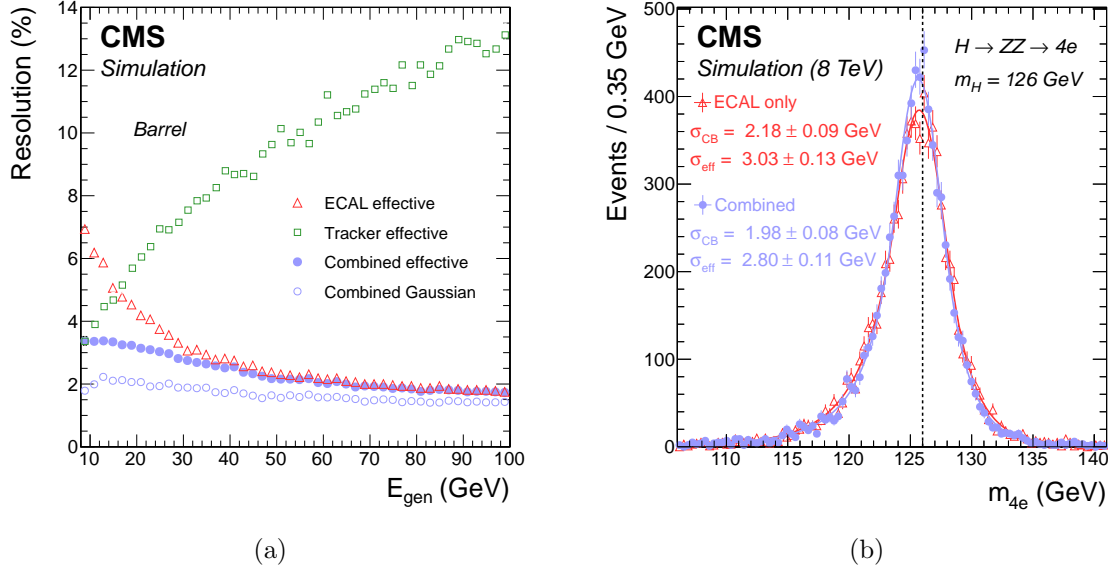


FIGURE 3.19 – (a) Résolution effective de l'impulsion des électrons après combinaison de E_{SC} et p (cercles pleins), comparée à celle de l'énergie SC corrigée (triangles), et de l'impulsion de la trace (carrés), en fonction de l'énergie générée pour des électrons dans le *barrel*. La résolution de l'impulsion combinée est également exposée en terme de σ_{CB} (cercles vides) pour illustrer la contribution du cœur gaussien de la distribution. Les électrons sont générés avec des distributions uniformes en η et p_{T} . (b) Masse invariante reconstruite du boson de Higgs pour des événements simulés $H \rightarrow ZZ^* \rightarrow 4e$ avec $m_H = 126$ GeV, en utilisant l'énergie SC corrigée (triangles vides) et l'impulsion après combinaison (cercles pleins) [34].

premier lieu pour l'impulsion correspondant à la combinaison, et en second lieu pour une gamme en p_{T} plus étendue, en particulier pour l'analyse du boson de Higgs en quatre leptons qui utilise des électrons avec des p_{T} aussi bas que 7 GeV. À cet effet, les événements $Z \rightarrow e^+e^-$ sont utilisés, complétés par les événements $J/\psi \rightarrow e^+e^-$ et $\Upsilon \rightarrow e^+e^-$, sources d'électrons de bas p_{T} . Les masses invariantes reconstruites de ces résonances sont comparées entre les données et la simulation pour accéder aux différences résiduelles.

La Fig. 3.20 montre deux exemples d'une telle comparaison et leur degré d'adéquation pour deux catégories extrêmes d'événements : une première dont les deux électrons sont bien mesurés, avec un SC contenant un *cluster* unique (*golden* ou *big-brem*) dans le *barrel*, et la seconde dont les deux électrons ont un SC contenant plusieurs *clusters* ou présentent un défaut de mesure (*showering*, *crack* ou *bad-track*) dans les *endcaps*. Les distributions dans les données et la simulation sont ajustées avec une fonction Breit–Wigner convoluée avec une fonction Crystal Ball,

$$P(m_{e^+e^-}; m_Z, \Gamma_Z, \alpha, n, m_{\text{CB}}, \sigma_{\text{CB}}) = \text{BW}(m_{e^+e^-}; m_Z, \Gamma_Z) \otimes f_{\text{CB}}(m_{e^+e^-}; \alpha, n, m_{\text{CB}}, \sigma_{\text{CB}}),$$

où m_Z et Γ_Z sont fixés aux valeurs nominales égales à 91.188 et 2.485 GeV [69].

La déviation standard effective σ_{eff} , indiquée sur les figures, est calculée comme la déviation standard effective de la fonction f_{CB} , ce qui de fait n'inclut pas la contribution venant de la largeur du boson Z. Les deux catégories montrent un bon accord entre les données et la simulation. Dans les données, les valeurs de σ_{eff} pour les masses invariantes de $Z \rightarrow e^+e^-$ sont, respectivement pour la première et seconde catégories, 1.13 ± 0.01 et 2.88 ± 0.02 GeV, les valeurs du cœur gaussien σ_{CB} sont 1.00 ± 0.01 et 2.63 ± 0.02 GeV. En conséquence, les résolutions effectives (gaussiennes) pour les masses invariantes de $Z \rightarrow e^+e^-$ sont 1.2 (1.1) %, pour la première catégorie, et 3.2 (2.9) %, pour la seconde. Les résolutions effectives (gaussiennes) par électron, approximées en multipliant par $\sqrt{2}$ les résolutions pour la masse sont donc comprises entre 1.7 (1.6) %, et 4.5 (4.1) %.

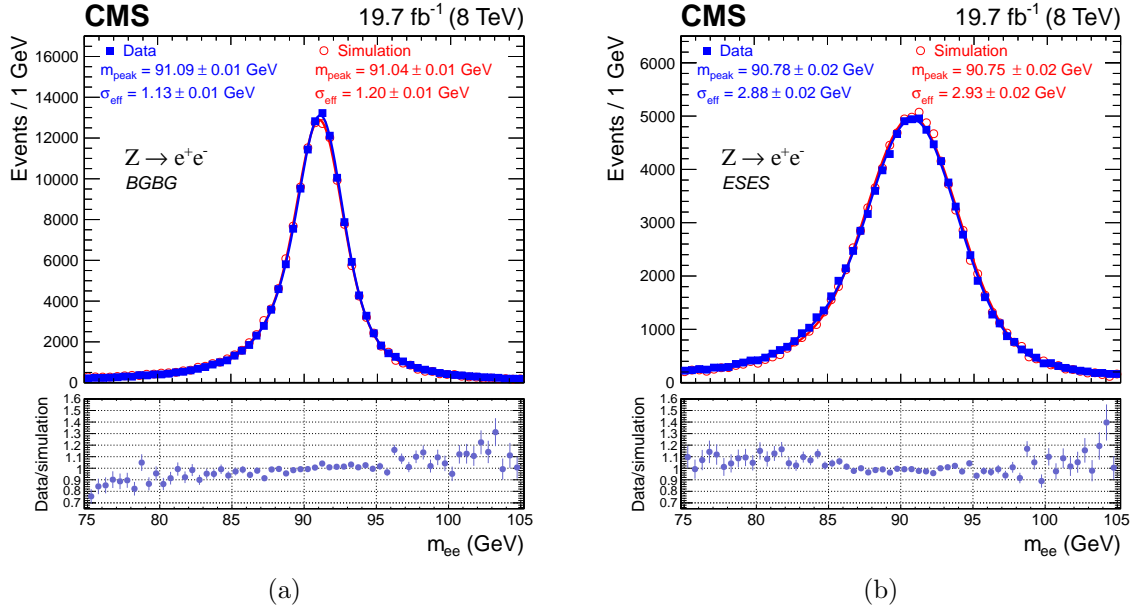


FIGURE 3.20 – Distributions de la masse invariante diélectron pour des événements de $Z \rightarrow e^+e^-$ dans les données (carrés pleins), comparées à la simulation (cercles vides), ajustées avec la convolution d’une fonction Breit–Wigner et d’une fonction Crystal Ball : (a) pour la meilleure catégorie avec deux électrons bien mesurés, dont le SC contient un *cluster* unique, dans le *barrel* (BGBG), et (b) pour la seconde avec deux électrons dont le SC contient plusieurs *clusters*, ou présentant un défaut de mesure, dans les *endcaps* (ESES). Les masses correspondant à la valeur maximale des fonctions d’ajustement m_{peak} et les déviations standards effectives σ_{eff} sont indiquées dans les figures. Les rapports entre les données et la simulation sont montrés en-dessous des figures.

La Table 3.4 montre les résolutions effectives et gaussiennes telles que définies précédemment (en pourcents) obtenues pour les masses invariantes de $Z \rightarrow e^+e^-$ dans les données, ainsi que les déviations standards effectives totales σ_{eff}^Z (incluant la largeur du Z), pour un ensemble plus complet de catégories, par ordre de qualité décroissante de résolution. Plus la résolution se dégrade, plus la distribution de masse s'éloigne d'une forme gaussienne.

Catégorie	$\sigma_{\text{eff}}/m_{\text{peak}}$ (%)	$\sigma_{\text{CB}}/m_{\text{peak}}$ (%)	$\sigma_{\text{eff}}^Z/m_{\text{peak}}$ (%)
BGBG	1.24	1.10	2.93
BGBS	1.81	1.55	3.41
BGEG	2.23	2.01	3.38
BSBS	2.30	2.01	3.84
EGEG	2.44	2.24	3.83
EGES	2.80	2.54	4.18
BSES	3.09	2.79	4.52
ESES	3.17	2.90	4.53
BB	1.92	1.64	3.53
EE	2.93	2.66	4.30

TABLE 3.4 – Résolutions effectives ($\sigma_{\text{eff}}/m_{\text{peak}}$) et gaussiennes ($\sigma_{\text{CB}}/m_{\text{peak}}$), en pourcents, obtenues pour les masses invariantes de $Z \rightarrow e^+e^-$ dans les données, et déviations standards effectives totales ($\sigma_{\text{eff}}^Z/m_{\text{peak}}$), pour différentes catégories. Pour la dénomination des catégories, B dénote un électron dans le *barrel* et E dans les *endcaps*, G désigne un électron bien mesuré dont le SC contient un *cluster* unique, S un électron dont le SC contient plusieurs *clusters*, ou présentant un défaut de mesure. Seules les catégories diagonales ou semi-diagonales sont montrées (les deux électrons sont au moins dans la même catégorie en η ou en classe).

Les comparaisons entre les données et la simulation sont effectuées pour différentes catégories d'événements, basées sur η , p_T , et la classe des électrons, pour différentes luminosités instantanées. Les corrections d'échelles sont appliquées sur les données, et les résolutions sont dégradées, comme discuté dans la Section 3.4.2.3.

Concernant l'étude sur l'échelle en impulsion, les catégories en p_T et η sont définies par rapport au p_T et η de l'un des électrons, l'autre électron étant utilisé pour étiqueter l'événement Z (comme décrit dans la Section 3.6) et vérifiant $p_T > 20$ GeV. Les ajustements sont réalisés en utilisant des modèles issus de la simulation (distributions discrètes) convolués avec des fonctions gaussiennes. Une dépendance de l'échelle en impulsion en fonction de p_T , valant jusqu'à 0.6 % dans le *barrel* et 1.5 % dans les *endcaps*, est observée et corrigée dans la gamme de 7 à 70 GeV. Le résultat final de l'échelle en impulsion est montré dans la Fig. 3.21 (a), représenté en tant que différence relative entre les données et la simulation des valeurs maximales des masses pour les événements $J/\psi \rightarrow e^+e^-$, $\Upsilon \rightarrow e^+e^-$, et $Z \rightarrow e^+e^-$, en

fonction du p_T et pour différentes régions en η d'un des électrons, en intégrant sur p_T et η de l'autre électron. La différence résiduelle d'échelle entre les données et la simulation atteint au plus 0.2 % dans le *barrel* et 0.3 % dans les *endcaps*. Ces nombres sont utilisés comme incertitudes systématiques pour l'échelle en impulsion des électrons dans le *barrel* et les *endcaps*.

Concernant l'étude sur la résolution, les catégories en p_T , η , et classe des électrons sont définies pour les deux électrons considérés. Les ajustements sont réalisés à l'aide d'une fonction Breit–Wigner convoluée avec une fonction Crystal Ball. L'accord entre les données et la simulation pour la résolution effective est montré dans la Fig. 3.21 (b), représenté en tant que différence relative entre les données et la simulation pour des événements $J/\psi \rightarrow e^+e^-$ et $Z \rightarrow e^+e^-$, en fonction du p_T d'un des électrons, pour différentes catégories d'électrons. Dans l'ensemble la différence relative entre les données et la simulation est inférieure à 10 % pour toutes les catégories considérées.

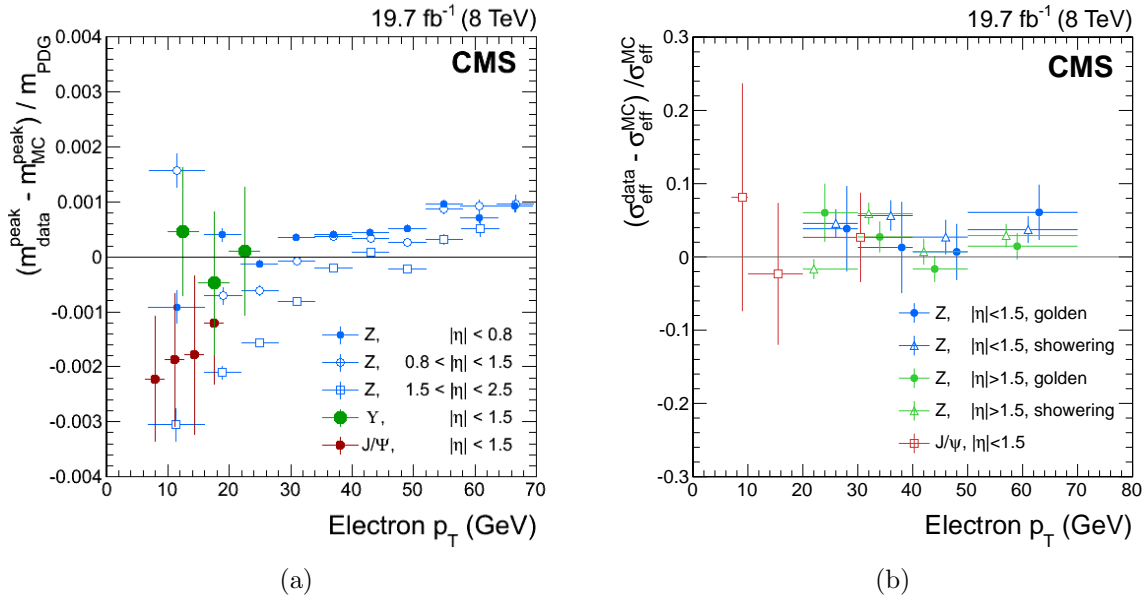


FIGURE 3.21 – Différences relatives entre les données et la simulation en fonction du p_T d'un des électrons et pour différentes régions en η : (a) pour l'échelle d'impulsion mesurée avec des événements $J/\psi \rightarrow e^+e^-$, $\Upsilon \rightarrow e^+e^-$, et $Z \rightarrow e^+e^-$ [34], et (b) pour la résolution effective en impulsion avec des événements $Z \rightarrow e^+e^-$ et $J/\psi \rightarrow e^+e^-$ pour différentes catégories d'électrons.

3.5 Sélection des électrons

3.5.1 Identification

Différentes stratégies sont exploitées dans CMS pour identifier les électrons primaires isolés (le signal) et les discerner des bruits de fond, venant principalement de conversions de photons, de jets faussement identifiés, ou de désintégrations semi-leptoniques des quarks b et c . Des algorithmes simples et robustes appliquant des sélections séquentielles sur un jeu d'observables discriminantes ont été développés. D'autres plus complexes combinent les observables dans une analyse multivariée afin d'obtenir une meilleure séparation.

Les observables discriminantes appartiennent principalement aux trois catégories suivantes :

- Les observables comparant les mesures obtenues par le ECAL et par le trajectographe (adéquation géométrique et cinématique entre la trace et le SC par exemple).
- Les observables purement calorimétriques, utilisées pour séparer les véritables électrons (électrons de signal ou issus des conversions de photons) des électrons faussement identifiés (dans les jets par exemple), sont fondées sur la forme transverse des cascades électromagnétiques dans le ECAL, et exploitent le fait que les cascades électromagnétiques sont plus étroites que les cascades hadroniques. Les énergies déposées dans le HCAL (qui doivent être petites, les cascades électromagnétiques étant essentiellement entièrement contenues dans le ECAL), ainsi que dans le *preshower* pour les *endcaps* sont également exploitées.
- Les observables correspondant aux traces améliorent la séparation entre les électrons et les hadrons chargés, en exploitant les informations contenues dans les traces ajustées par l'algorithme GSF, et les différences entre les informations provenant des ajustements GSF et KF.

Un exemple d'observable de trace, f_{brem} , a été exposé dans la Fig. 3.9. La Fig. 3.22 montre des exemples d'observables purement calorimétriques et d'association entre trace et SC. Le signal simulé correspond à des électrons reconstruits compatibles avec ceux générés dans des événements $Z \rightarrow e^+e^-$, en utilisant une version de la simulation dépendante de la période de prise de données. Les données correspondent à des électrons reconstruits dans un échantillon dominé par des événements $Z \rightarrow e^+e^-$. Afin que la pureté soit suffisante dans les données, un critère sévère sur la masse invariante des deux électrons est appliqué dans les données ainsi que dans la simulation : $|m_{e^+e^-} - m_Z| < 7.5 \text{ GeV}$. Les deux électrons doivent également être isolés : la somme scalaire des impulsions transverses des candidats PF dans un cône autour de la direction de chaque électron (excluant l'électron) ne doit pas excéder 10 % de son p_T . Le bruit de fond consiste en électrons faussement identifiés provenant de jets dans des données Z+jets. Les événements de cet échantillon sont sélectionnés en requérant une paire de leptons identifiés (électrons ou muons) formant une masse invariante compatible avec celle du Z. Afin de supprimer des contributions provenant de production associée de bosons W et Z, l'énergie transverse manquante dans les événements doit être inférieure à

25 GeV (ce qui supprime également des événements $t\bar{t}$). Un candidat électron additionnel doit être présent dans chaque événement considéré, pour lequel le critère d'isolation est inversé. Dans les événements $e^+e^- + \text{jets}$, la masse invariante entre le candidat électron faussement identifié et l'électron de charge opposée provenant du $Z \rightarrow e^+e^-$ est requise supérieure à 4 GeV, afin de rejeter les contributions venant des résonances à très basse masse invariante. En conséquence, l'échantillon de bruit de fond consiste principalement en événements avec un boson Z et un jet faussement identifié en électron. Tous les électrons de signal et de fond sont sélectionnés avec $p_T > 20$ GeV.

La distance en η entre la trace et le SC, définie précédemment (Section 3.3.4), est montrée dans les Figs. 3.22 (a) et (b). Alors que l'accord entre les données et la simulation est très bon dans le *barrel*, un désaccord est observé dans les *endcaps*, relié à la mauvaise modélisation du matériel dans la simulation. En effet, $|\Delta\eta|$ augmente avec la quantité de bremsstrahlung, qui dans les *endcaps* est plus importante pour les données que pour la simulation.

L'extension latérale de la cascade électromagnétique suivant η est exprimée par l'observable $\sigma_{\eta\eta}$, définie comme $(\sigma_{\eta\eta})^2 = \frac{\sum (\eta_i - \bar{\eta})^2 w_i}{\sum w_i}$. La somme concerne la matrice de 5×5 cristaux autour du cristal de plus haut E_T du SC, et w_i est un poids qui dépend logarithmiquement de l'énergie contenue dans chaque cristal. Les positions η_i sont exprimées en unités de cristaux, ce qui permet de prendre en compte simplement la taille variable des interstices non instrumentés entre les cristaux (en particulier aux frontières entre les modules). Cette observable est montrée dans les Figs. 3.22 (c) et (d). Le pouvoir de discrimination de $\sigma_{\eta\eta}$ est plus important que pour l'observable analogue en ϕ , la structure en ϕ du dépôt d'énergie dans le ECAL étant très perturbée par la présence de bremsstrahlung. Un léger désaccord est observé entre les données et la simulation dans le *barrel*, principalement dû à la modélisation limitée des cascades électromagnétiques dans la simulation (amélioré dans la version 10.0 de GEANT4 [70]). Pour les électrons dans les *endcaps*, le facteur déterminant dans la résolution des observables de forme de cascade étant l'empilement, bien décrit par la version utilisée de la simulation, l'accord entre les données et la simulation est bon dans ces figures.

Enfin, les Figs. 3.22 (e) et (f) montrent les distributions de $1/E_{SC} - 1/p$, où E_{SC} est l'énergie du SC et p l'impulsion de la trace au point d'approche minimale du vertex. Un bon accord est observé entre les données et la simulation pour le *barrel* et les *endcaps*. Dans tous les cas, la discrimination entre le signal et le bruit de fond est significative.

Les sélections séquentielles sont appliquées sur cinq observables d'identification : $|\Delta\eta|$, $\sigma_{\eta\eta}$, $1/E_{SC} - 1/p$ (montrées dans la Fig. 3.22), ainsi que $|\Delta\phi|$ et H/E_{SC} (définies précédemment, dans la Section 3.3.4 et 3.3.2.1, respectivement). Quatre niveaux de sévérité sont définis : *veto*, *loose*, *medium* et *tight*, par ordre croissant de sévérité.

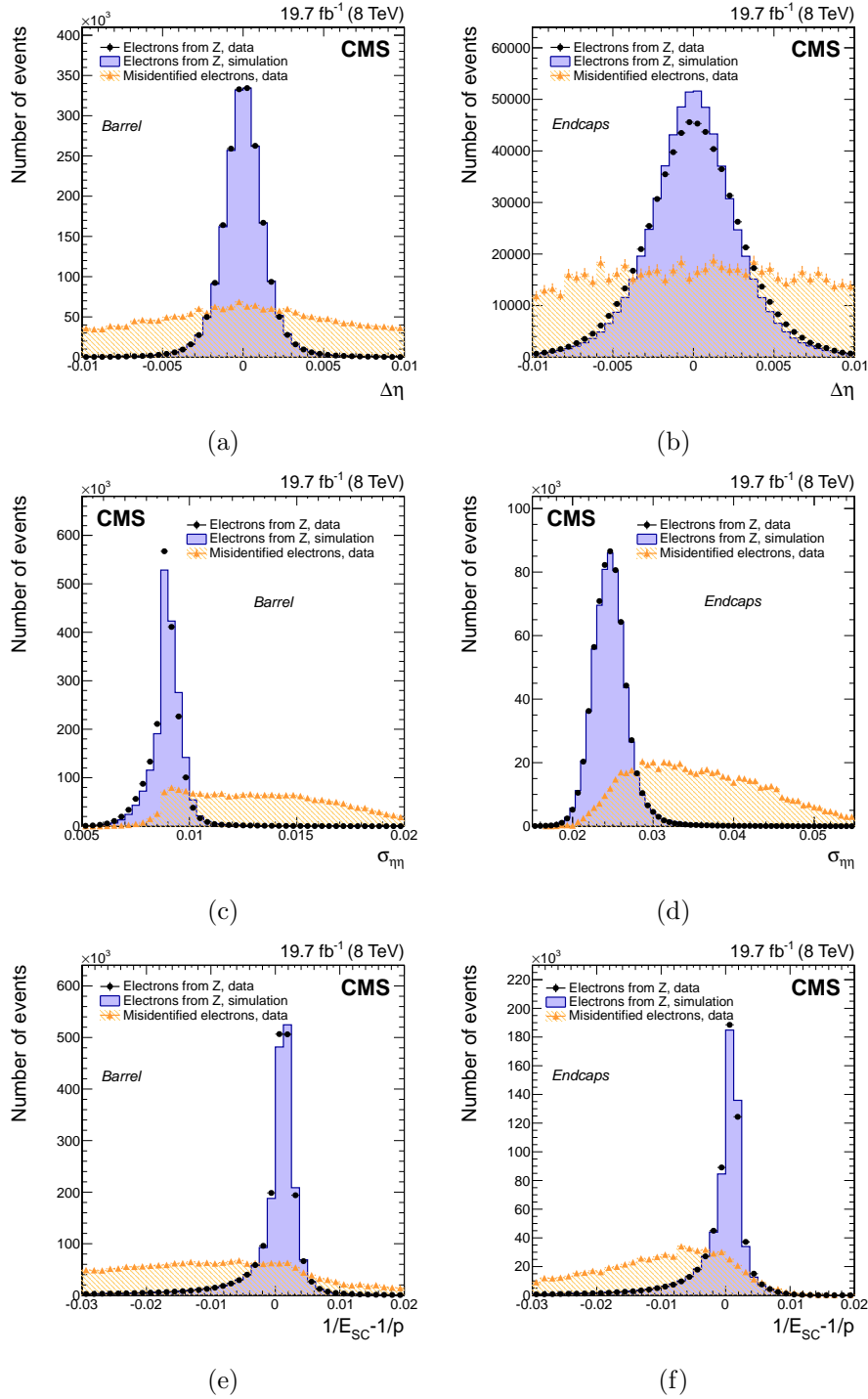


FIGURE 3.22 – Distributions de : (a) et (b) la distance $\Delta\eta$ entre la position du SC et celle de la trace, (c) et (d) l’observable de forme de cascade $\sigma_{\eta\eta}$, (e) et (f) l’association entre l’énergie du SC et l’impulsion de la trace $1/E_{SC} - 1/p$; pour les électrons des événements $Z \rightarrow e^+e^-$ dans les données (points) et la simulation (histogrammes pleins), et pour des électrons faussement identifiés dans les données (triangles); (a) (c) (e) dans le *barrel*, et (b) (d) (f) dans les *endcaps*. Chaque distribution est normalisée à l’aire correspondante dans les événements de données $Z \rightarrow e^+e^-$.

Afin de maximiser la sensibilité de l’identification des électrons, plusieurs observables sont combinées en utilisant un algorithme multivarié de *boosted decision tree* (BDT) [55]. L’ensemble des variables de chaque catégorie est étendu par rapport à l’approche de sélection séquentielle. En particulier, les observables d’association entre la trace et le SC sont calculées à la fois à la surface du ECAL et au vertex, la structure du SC est mieux exploitée en utilisant davantage d’observables de forme de cascade, et la fraction f_{brem} est utilisée. Deux types de BDT sont définis, dépendant si l’électron considéré a déclenché un *trigger* HLT (“électron *triggering*”) ou non (“électron *not-triggering*”). Pour les électrons *triggering*, une présélection est appliquée correspondant aux critères d’identification et d’isolation appliqués au HLT, ainsi qu’un entraînement dédié qui permet d’exploiter au mieux le pouvoir discriminant des observables dans l’espace de phase résiduel. Le type *not-triggering* est exploité dans les topologies avec plusieurs leptons, en particulier pour l’analyse du boson de Higgs en quatre leptons, pour laquelle seule une partie des leptons participe au déclenchement.

Par la suite, les résultats sont présentés uniquement pour les électrons *not-triggering*. Le BDT est entraîné dans différents intervalles en p_T et η . Les électrons reconstruits et associés avec ceux générés sont utilisés pour modéliser le signal, dans une gamme en p_T entre 5 et 100 GeV. Le bruit de fond est modélisé grâce à des électrons faussement identifiés dans des événements W+jets dans les données. Les distributions des observables discriminantes dans ces échantillons d’entraînement sont en accord avec celles dans les échantillons d’analyse où l’identification est appliquée. Les distributions de sortie du BDT pour le signal et le bruit de fond sont comparées dans la Fig. 3.23, avec en outre pour le signal une comparaison entre les données et la simulation. Les mêmes échantillons et sélections que pour la Fig. 3.22 sont utilisés pour le signal (dans les données et la simulation) et le bruit de fond. Le pouvoir discriminant de l’algorithme BDT est indéniable, et l’accord entre les données et la simulation est bon. Les quelques petites différences observées sont reliées aux différences observées dans les observables d’entrée, décrites dans les paragraphes précédents. Elles sont prises en compte par des corrections apportées aux efficacités d’identification extraites de l’analyse des événements $Z \rightarrow e^+e^-$, comme discuté dans la Section 3.6.

Les performances de l’identification BDT sont comparées avec celles de la sélection séquentielle pour quatre points de fonctionnement dans la Fig. 3.24 pour des électrons avec $p_T > 20$ GeV. Les électrons de signal dans des événements $Z \rightarrow e^+e^-$ simulés sont comparés avec des électrons faussement identifiés dans des jets issus des données. Les mêmes échantillons et sélections que pour la Fig. 3.22 sont utilisés. Comme attendu, les performances obtenues sont meilleures quand les observables sont combinées dans un discriminant multivarié tel que le BDT. Dans le *barrel* et les *endcaps* respectivement, un point de fonctionnement de la sélection séquentielle avec une efficacité pour les électrons de signal d’approximativement 90 et 84 % présente une efficacité d’approximativement 7 et 9 % sur le bruit de fond. Avec la même efficacité sur le signal, l’efficacité de fausse identification sur le bruit de fond est réduite d’un facteur deux par l’identification BDT.

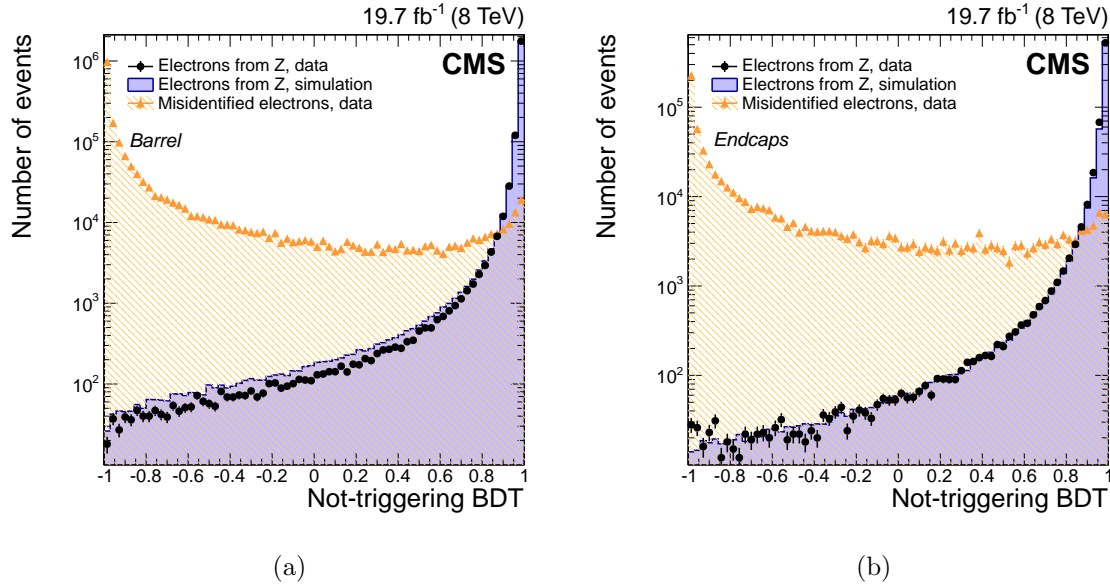


FIGURE 3.23 – Observable de sortie du BDT de l’identification des électrons pour les électrons provenant d’événements $Z \rightarrow e^+e^-$ dans les données (points) et la simulation (histogrammes pleins), et pour des électrons faussement identifiés dans les données (triangles), (a) dans le *barrel*, et (b) dans les *endcaps*. Chaque distribution est normalisée à l’aire correspondante dans les événements de données $Z \rightarrow e^+e^-$.

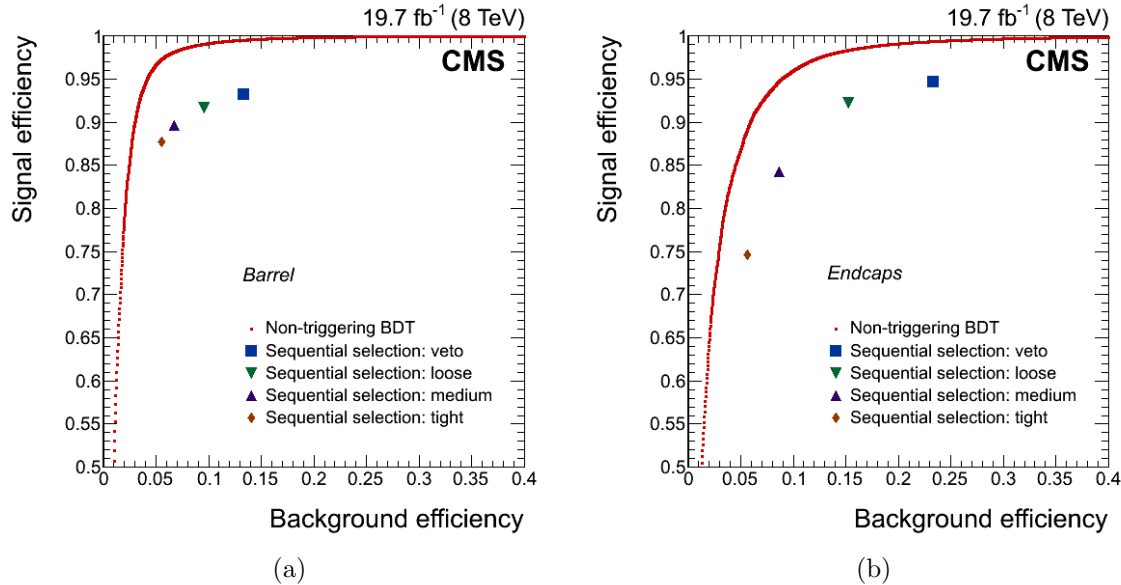


FIGURE 3.24 – Performances de l’identification BDT (points rouges), comparées à celles des points de fonctionnement de la sélection séquentielle pour des électrons dans (a) le *barrel*, et (b) les *endcaps*.

Le lecteur attentif aura remarqué que les résultats illustrés jusqu'à présent sur l'identification concernent des électrons avec $p_T > 20$ GeV. Cette sélection est choisie uniquement dans un but illustratif, afin d'avoir un échantillon de signal suffisamment pur pour que les distributions ne soient pas biaisées. La même stratégie d'identification est également performante et bien sûr utilisée pour des p_T inférieurs, bien que le contrôle des performances et la validation des algorithmes soient moins aisés pour ceux-ci. L'accord entre les données et la simulation a été étudiée dans la gamme $7 < p_T < 15$ GeV grâce au méson J/ψ . La Fig. 3.25 illustre cette comparaison pour deux observables, pour des événements avec deux électrons dans le *barrel* et une simulation dépendante de la période de prise de données. Le bruit de fond est soustrait statistiquement, en utilisant la technique *sPlot* [71] et un ajustement de la masse invariante diélectron. L'accord entre les données et la simulation est bon pour les variables telles que $\sigma_{\eta\eta}$ dans la Fig. 3.25 (a), tout comme pour des observables plus complexes tels que la sortie du BDT montrée dans la Fig. 3.25 (b).

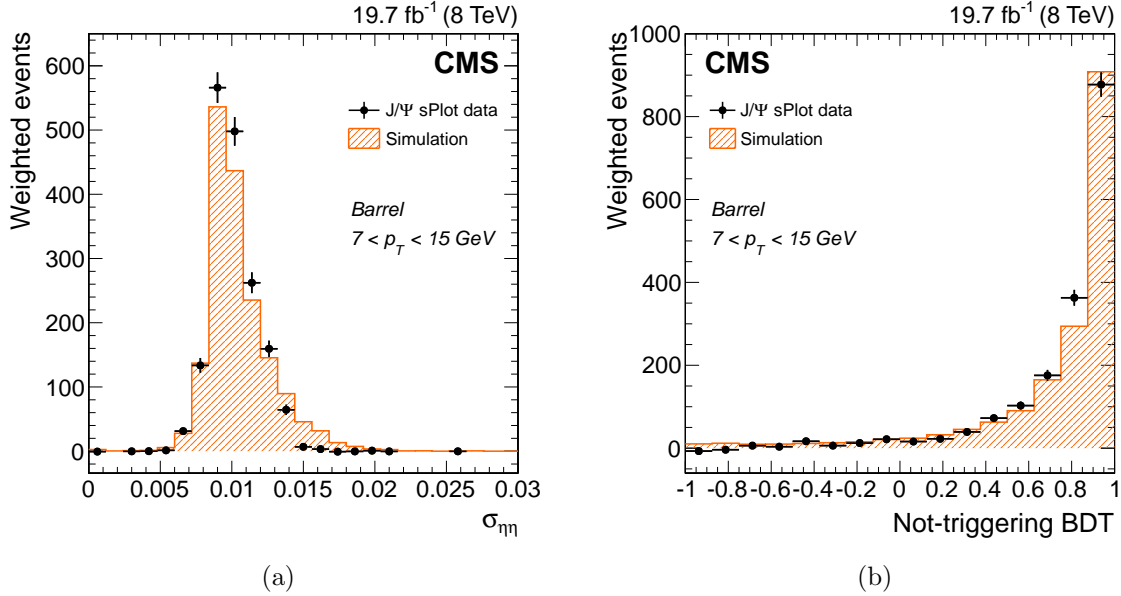


FIGURE 3.25 – Distributions de (a) l'observable de forme de cascade $\sigma_{\eta\eta}$, et (b) la sortie d'identification BDT pour des électrons provenant de J/ψ , avec $7 < p_T < 15$ GeV dans le *barrel*, pour les données (symboles) et la simulation (histogrammes). Une soustraction statistique du bruit de fond est appliquée en suivant la technique *sPlot*.

3.5.2 Isolation

Une fraction significative du bruit de fond pour les électrons primaires isolés provient de jets faussement identifiés ou d'électrons véritables résultant de désintégrations semi-leptoniques de quarks b ou c dans un jet. Dans ces cas, un flux d'énergie important est présent autour de la trajectoire des candidats électrons, et l'application d'un critère d'isolation par rapport à une telle activité réduit considérablement ces sources de bruit de fond. L'isolation est distincte de l'identification, la balance entre les deux dépendant fortement de l'analyse considérée ; de plus, l'inversion du critère d'isolation indépendamment de l'identification (et inversement) permet le contrôle de différentes sources de bruit de fond dans les données.

Deux techniques sont utilisées dans CMS pour l'isolation. La plus simple est basée sur les détecteurs de façon distincte et s'appuie sur la somme des énergies déposées dans le ECAL et le HCAL, et sur la somme scalaire des p_T des traces reconstruites depuis le vertex de la collision, à l'intérieur d'un cône de rayon $\Delta R = 0.3$ ou 0.4 autour de la trajectoire de l'électron, tout en soustrayant la contribution de l'électron considéré dans des cônes de tailles inférieures. Cette procédure est utilisée dans le HLT, par certaines analyses ne nécessitant pas les performances d'isolation optimales, et pour les électrons de très hautes énergies. Cependant, la plupart des analyses bénéficient de la technique PF, définissant dans un cône d'isolation les particules elles-mêmes plutôt que les mesures distinctes dans chaque détecteur. Ainsi, une calibration correcte est utilisée, les doubles comptages éventuels sont évités, et la contribution de l'électron considéré est naturellement exclue du calcul de l'observable d'isolation. Cependant, si l'électron considéré est faussement identifié comme une autre particule par le PF (un hadron chargé par exemple), il entre dans le calcul de l'observable d'isolation et l'accroît artificiellement. Cet effet est d'autant plus important que l'efficacité d'identification de l'électron par le PF diminue. Négligeable dans le *barrel*, il devient significatif dans les *endcaps* (du moins pour les échantillons de données et les versions des algorithmes discutés dans ce document) et nécessite une correction supplémentaire qui s'appuie sur des cônes d'exclusions comme dans le cas de l'isolation basée sur les détecteurs.

Une comparaison entre les performances des deux techniques est montrée dans la Fig. 3.26 pour des électrons avec $p_T > 20$ GeV (sans application de correction pour l'empilement). Les électrons de signal dans des événements $Z \rightarrow e^+e^-$ simulés sont comparés avec des électrons faussement identifiés dans des jets de données. Une version de la simulation dépendante de la période de prise de données est utilisée. Comme attendu, une meilleure performance est obtenue quand les informations concernant tous les détecteurs sont combinées par la technique PF, en particulier dans les *endcaps*.

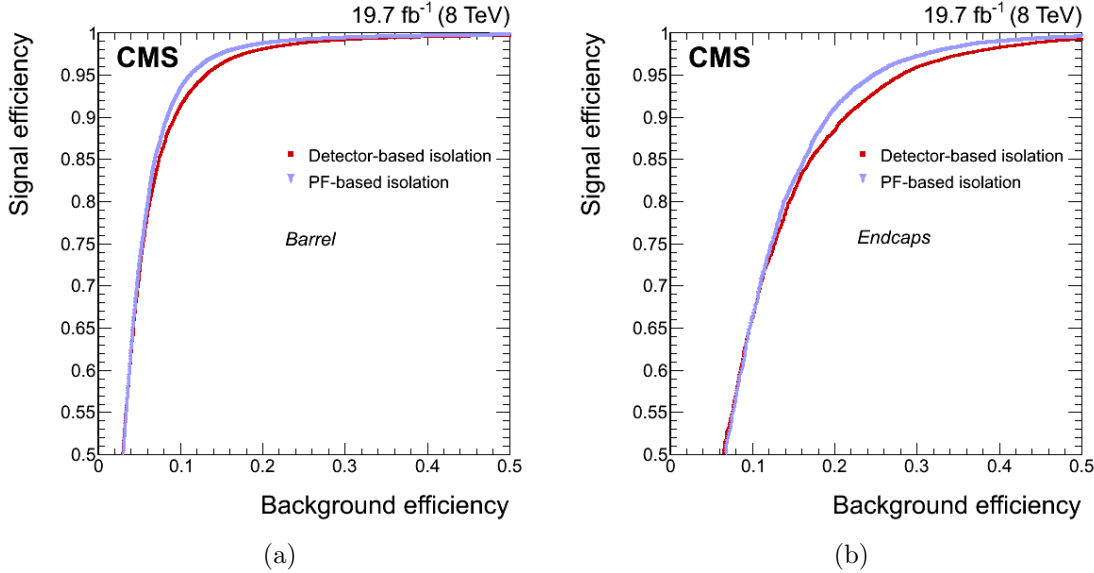


FIGURE 3.26 – Performances de l’isolation basée sur les détecteurs de façon distincte (carrés rouges) comparées à celles de l’isolation PF (triangles bleus) dans (a) le *barrel*, et (b) les *endcaps*.

L’isolation PF des électrons est définie comme

$$\text{ISO}_{\text{PF}} = \sum p_{\text{T}}^{\text{charged}} + \text{Max}[0, \sum p_{\text{T}}^{\text{neutral had}} + \sum p_{\text{T}}^{\gamma} - p_{\text{T}}^{\text{PU}}], \quad (3.2)$$

où la somme concerne toutes les particules PF chargées (*charged*), les hadrons neutres (*neutral had*) et les photons (γ), à l’intérieur d’un cône ΔR autour de la direction de l’électron. Les particules chargées ont leur origine au vertex de l’événement considéré, et p_{T}^{PU} est la correction relative à l’empilement.

En effet, les observables reliées à l’isolation sont parmi celles les plus sensibles à l’énergie supplémentaire provenant d’interactions d’empilement (pendant le croisement de faisceaux ou provenant de croisements antérieurs), ce qui dégrade les performances, et en particulier l’efficacité de sélection. La contribution provenant de l’empilement à l’intérieur du cône d’isolation, qui doit par conséquent être soustraite de l’observable d’isolation, est estimée avec la technique FASTJET [72–74], qui suppose $p_{\text{T}}^{\text{PU}} = \rho A_{\text{eff}}$. La densité ρ (définie dans la Section 3.4.2.2) dépend de l’empilement comme illustré dans la Fig. 3.27 (a) pour des événements de données $Z \rightarrow e^+e^-$, le nombre d’événements d’empilement étant relié au nombre de vertex de collisions dans l’événement. La dépendance des composantes chargée et neutre de l’isolation PF y est également représentée. La composante chargée est indépendante de l’empilement, les particules chargées considérées provenant uniquement du vertex de l’événement. En revanche, ρ et la composante neutre dépendent quasi-linéairement du nombre de vertex de collisions dans l’événement. La surface effective A_{eff} en (η, ϕ) est définie, pour chaque

composante de l'isolation, par la valeur de $(\Delta R)^2$ pondérée par le rapport entre les pentes de ρ et la composante considérée montrée dans Fig. 3.27 (a). Une fois la correction appliquée aux composantes neutres, la dépendance en fonction du nombre de vertex est considérablement réduite, comme illustré dans la Fig. 3.27 (b). Les figures sont montrées pour des électrons avec $|\eta| < 1$, mais les conclusions sont les mêmes pour toute la gamme en η .

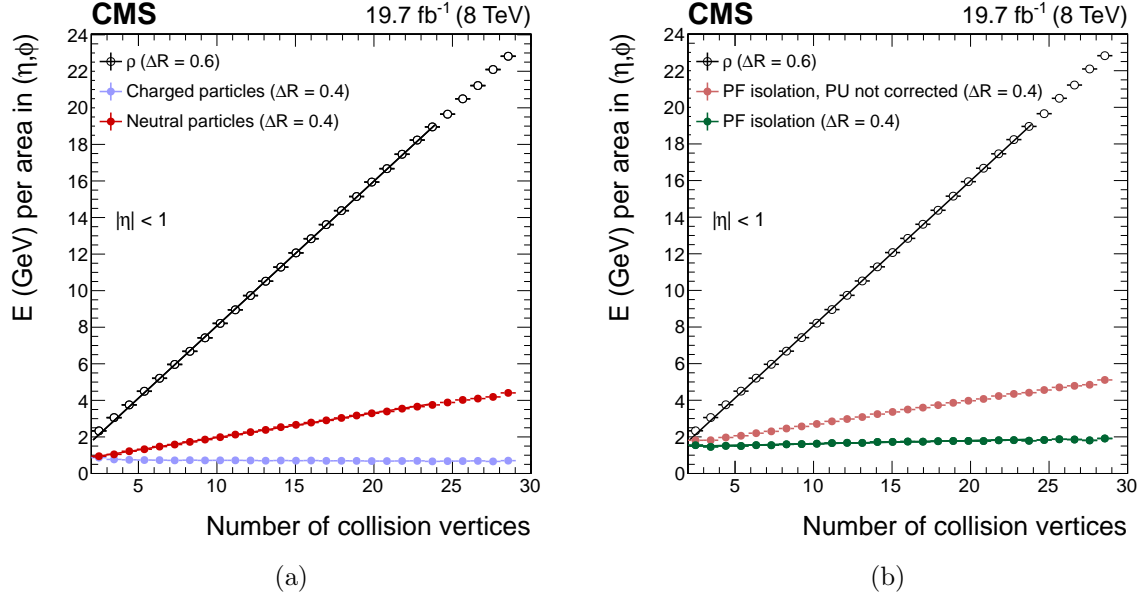


FIGURE 3.27 – Densité d'énergie moyenne en fonction du nombre de vertex de collisions reconstruits, pour des électrons avec $p_T > 20$ GeV et $|\eta| < 1$ dans des événements $Z \rightarrow e^+e^-$ dans les données : (a) ρ (cercles noirs), particules chargées provenant du vertex de l'événement (points bleus) et particules neutres (points rouges), et (b) isolation PF totale, avant (points roses) et après (points verts) la correction d'empilement.

La Fig. 3.28 présente les distributions de l'observable d'isolation Iso_{PF} divisée par le p_T de l'électron, pour des électrons de signal (dans les données et la simulation) et des électrons faussement identifiés dans les données, après la correction concernant les contributions d'empilement. Les échantillons et sélections sont ceux présentés dans la Section 3.5.1, à l'exception du critère d'isolation qui est remplacé par une sélection relâchée sur l'identification BDT. Une excellente discrimination entre le signal et le bruit de fond est observée, ainsi qu'un bon accord entre les données et la simulation. Le désaccord résiduel dans les *endcaps* est due principalement à une différence de l'efficacité de l'identification PF entre les données et la simulation, qui implique une différence de la contribution de particules faussement identifiées dans l'observable d'isolation. Celle-ci n'est pas complètement corrigée par l'utilisation des cônes d'exclusion.

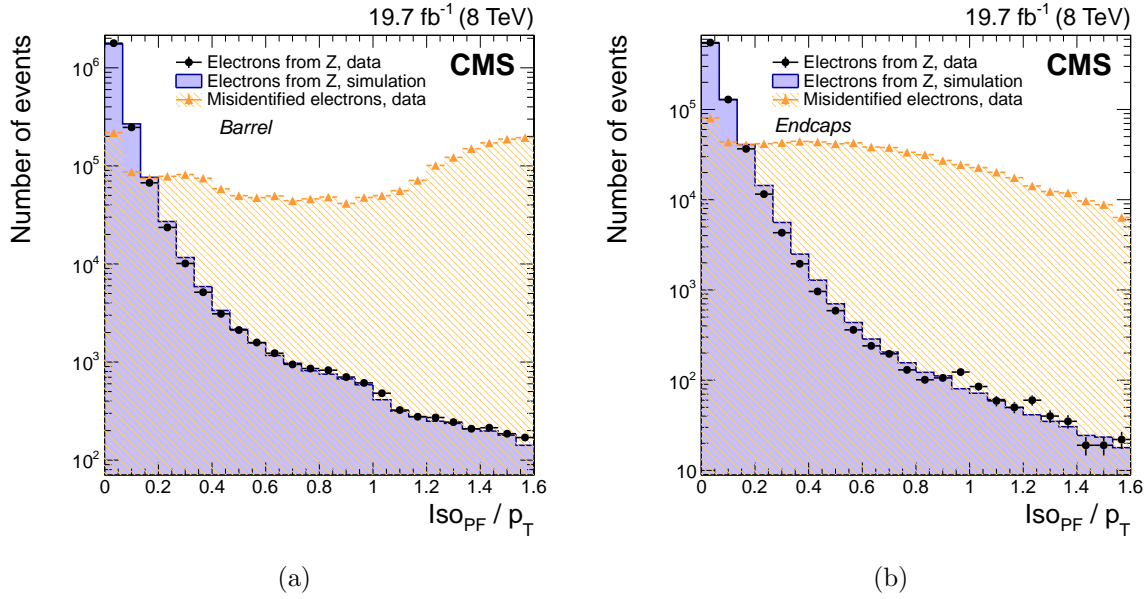


FIGURE 3.28 – Distributions de l’observable d’isolation PF divisée par le p_T de l’électron, après application de la correction d’empilement, pour les électrons provenant d’événements $Z \rightarrow e^+e^-$ dans les données (points) et la simulation (histogrammes pleins), et pour des électrons faussement identifiés dans les données (triangles), dans (a) le *barrel*, et (b) les *endcaps*.

3.5.3 Réjection des photons convertis

Une source importante de bruit de fond pour les électrons primaires provient des électrons secondaires produits par les conversions de photons dans le matériel du trajectographe. L’exploitation des schémas de *hits* des traces permet la réjection de ce bruit de fond. Les électrons reconstruits résultant d’une conversion de photon dans le trajectographe ont fréquemment leur premier *hit* distant de la couche la plus interne du trajectographe. Ainsi, des *hits* internes manquants (tels que définis dans la Section 3.3.5) peuvent être constatés dans cette région. À l’opposé, les électrons primaires sont issus directement du vertex, et aucun *hit* interne manquant n’est attendu dans les couches internes du trajectographe.

En complément des *hits* internes manquants, les électrons issus de conversions de photons peuvent être rejetés en utilisant un ajustement des traces des électrons. Le photon étant sans masse et le transfert d’impulsion généralement faible, les conversions possèdent une topologie caractéristique, avec des traces ayant essentiellement la même tangente au vertex de conversion, dans les plans (r, ϕ) et également (r, z) . La stratégie consiste donc à ajuster les paires de traces d’électrons avec leur vertex commun en incluant cette contrainte topologique, puis à rejeter les événements en fonction de la probabilité de χ^2 obtenue par l’ajustement.

Les paramètres d'impact des électrons sont également utilisés pour rejeter les électrons secondaires : les distances transverse (d_0) et longitudinale (d_z) entre le vertex et le point d'approche minimale dans le plan transverse, et le rapport entre le paramètre d'impact à trois dimensions et son incertitude (σ_{ip}/ip).

De façon générale, un critère requérant aucun *hit* interne manquant associé à une sélection sur la probabilité de χ^2 décrite ci-dessus présente une inefficacité de l'ordre du pourcent sur des électrons provenant d'événements $Z \rightarrow e^+e^-$ simulés, pour un facteur de réjection sur le bruit de fond d'approximativement 45 %. Ces nombres sont donnés à titre indicatif, car ils dépendent fortement des sélections préalablement appliquées sur les électrons, en modifiant la composition du bruit de fond et par conséquent la fraction des conversions de photons. Pour l'exemple cité ci-dessus, le critère d'identification correspondant aux électrons dans l'analyse du Higgs en quatre leptons (détaillé dans la section suivante) est appliqué, en omettant la sélection sur le nombre de *hits* internes manquants.

3.5.4 Sélection des électrons pour l'analyse $H \rightarrow ZZ^* \rightarrow 4\ell$

À chaque analyse de physique de trouver la bonne balance entre efficacité et pureté, dépendant des quantités de signal et de bruits de fond et de la finalité poursuivie, en définissant sa propre combinaison des différents algorithmes de sélection présentés. Cette section décrit la sélection des électrons dans l'analyse du boson de Higgs se désintégrant en quatre leptons [34].

Celle-ci combine l'identification BDT décrite dans la Section 3.5.1 (pour des électrons *not-triggering*), l'isolation PF, et les observables de réjection des électrons secondaires, et est détaillée dans la Table 3.5. Elle est optimisée dans le but de maximiser la sensibilité attendue pour un boson de Higgs de basse masse, pour deux catégories en p_T ($7 < p_T < 10$ GeV et $p_T > 10$ GeV) et trois en $|\eta|$ ($|\eta| < 0.80$, $0.80 < |\eta| < 1.48$, et $1.48 < |\eta| < 2.50$), correspondant à deux régions du *barrel* avec des quantités différentes de matériel devant le ECAL, et une région dans les *endcaps*.

Observable			$ \eta < 0.8$	$0.8 < \eta < 1.479$	$ \eta > 1.479$
$7 < p_T < 10$ GeV	BDT	>	0.47	0.004	0.295
$p_T > 10$ GeV	BDT	>	-0.34	-0.65	0.60
Iso _{PF} / p_T ($\Delta R = 0.4$)				0.40	
$ \sigma_{ip}/ip $				4	
<i>hits</i> internes manquants				1	

TABLE 3.5 – Sélection pour les électrons dans l'analyse $H \rightarrow ZZ^* \rightarrow 4\ell$ [34].

3.6 Mesures d’efficacité et de probabilité de fausse identification

Une méthode basée sur la technique de *tag-and-probe* (T&P) [37], très largement utilisée dans CMS et en physique de haute énergie auprès des collisionneurs, exploite les événements $Z/\gamma^* \rightarrow e^+e^-$ dans les données afin d’estimer les efficacités de reconstruction et de sélection pour les électrons de signal. La méthode sélectionne des événements $Z/\gamma^* \rightarrow e^+e^-$, pour lesquels un électron vérifie des critères sévères et sert d’identifiant (*tag*), alors que le deuxième électron est utilisé comme sonde (*probe*) pour la mesure de l’efficacité considérée.

Différents critères ont été testés pour la sélection de l’électron *tag* et les résultats montrent que l’estimation des efficacités dépend très peu de la définition du *tag*. Pour les résultats décrits par la suite, les électrons *tag* vérifient $p_T > 25$ GeV et sont identifiés par le point de fonctionnement le plus sévère de la sélection séquentielle (*tight*). L’électron *probe* permet de mesurer l’efficacité considérée et sa sélection dépend donc de l’algorithme étudié. La masse invariante des deux électrons doit vérifier $60 < m_{e^+e^-} < 120$ GeV, région suffisamment large pour permettre l’évaluation des bruits de fond par l’ajustement. Deux électrons de charges opposées sont également requis lorsque cela est rendu possible par la sélection de l’électron *probe*. Enfin, les événements utilisés par la procédure de T&P satisfont des *triggers* HLT n’induisant pas de biais pour l’efficacité étudiée. Ainsi, dans le cas de l’efficacité de reconstruction, les *triggers* utilisés sont ceux requérant un électron et un SC, avec l’électron *tag* associé à l’électron du HLT et l’électron *probe* au SC du HLT. Dans le cas de l’efficacité de sélection, les *triggers* utilisés sont ceux requérant deux électrons avec des sélections moins restrictives que celles à l’étude, et chacun des électrons *tag* et *probe* sont associés avec un électron du HLT. Toutes les paires *tag*–*probe* de l’événement sont considérées afin de réduire les biais éventuels introduits par un choix arbitraire.

Le nombre d’électrons *probe* passant une sélection est déterminé par deux ajustements simultanés de la distribution de la masse invariante, pour le cas où les électrons *probe* passent la sélection, et le cas inverse. Les ajustements incluent les contributions pour le signal et le bruit de fond. De façon générale le bruit de fond est décrit par une fonction exponentielle, convoluée avec une fonction d’erreur qui reflète les seuils cinématiques sur le p_T de l’électron *probe*. Tous les paramètres de l’exponentielle et de la fonction d’erreur sont laissés libres lors de l’ajustement. L’ajustement pour le signal utilise généralement des expressions analytiques, ou des modèles issus de la simulation (distributions discrètes). Les expressions analytiques sont composées d’une fonction Breit–Wigner avec la masse et la largeur naturelle du boson Z [69] convoluée avec une fonction Crystal Ball jouant le rôle de fonction de résolution, et multipliée par une exponentielle décroissante afin de modéliser le signal dans la région de masse invariante entre 60 et 70 GeV. Les modèles sont issus d’électrons dans des événements $Z \rightarrow e^+e^-$ simulés, convolués avec une fonction de résolution tenant compte des différences résiduelles entre les données et la simulation.

La même technique de T&P est appliquée aux données et à la simulation afin de comparer les efficacités et obtenir des rapports entre les données et la simulation. Pour la grande majorité des analyses, ceux-ci sont appliqués en tant que facteurs correctifs de la simulation ; pour les autres ils sont utilisés pour l'évaluation des incertitudes systématiques. Pour la simulation, les efficacités sont évaluées sur des échantillons de signal $Z \rightarrow e^+e^-$ sans addition de bruits de fond. Une association géométrique entre les électrons générés et reconstruits est souvent requise pour résoudre les ambiguïtés, en particulier à bas p_T . Les ajustements sont réalisés par intervalle de η et p_T .

La Fig. 3.29 illustre ces ajustements dans les données avec un exemple de sélection et d'intervalle en p_T et η , pour (a) le cas où les électrons *probe* passent la sélection, et (b) le cas inverse. Le signal observé dans la région de masse entre 60 et 70 GeV correspond aux contributions provenant d'événements γ^* , d'événements avec radiation dans l'état final, et d'événements contenant des électrons mal mesurés, essentiellement situés dans les régions non instrumentées du ECAL.

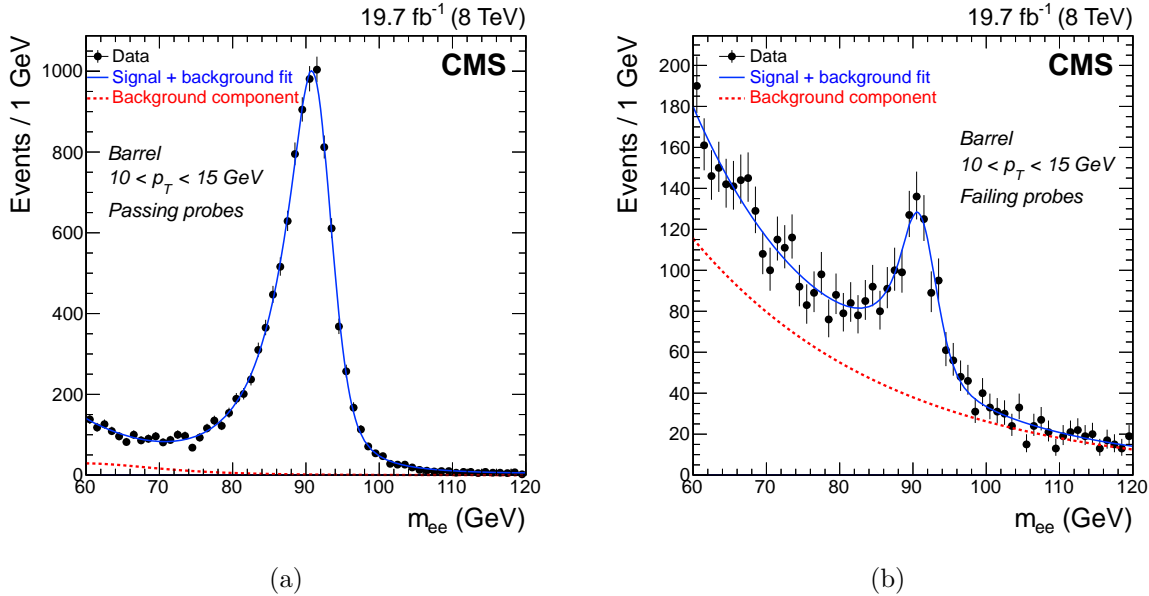


FIGURE 3.29 – Exemple d'ajustements des distributions de la masse invariante dans les données, pour des électrons *probe* avec $10 < p_T < 15$ GeV dans le *barrel* : (a) satisfaisant, et (b) ne satisfaisant pas, les sélections sur l'isolation et le paramètre d'impact de la sélection pour l'analyse du boson de Higgs en quatre leptons [34]. Les ajustements sont montrés pour l'ensemble signal + bruits de fond (ligne continue) et pour le bruit de fond (ligne en pointillé).

Les principales incertitudes systématiques proviennent du choix des fonctions utilisées pour les ajustements et sont estimées par la comparaison des résultats issus de choix alternatifs pour les fonctions et modèles de signal et de bruit de fond. Les autres sources

d'incertitudes systématiques, telles que le choix du nombre d'intervalles pour chaque ajustement, la définition de l'électron *tag* et le type de générateur utilisé pour la simulation ont montré des influences négligeables.

3.6.1 Efficacité de reconstruction

L'efficacité de reconstruction est évaluée en fonction de E_T^{SC} et η_{SC} , et concerne l'efficacité de la reconstruction de la trace et de l'association avec le SC, l'efficacité de reconstruction des SC pour $E_T^{SC} > 5$ GeV étant très proche de 100 %. Les mesures de l'efficacité de la reconstruction dans les données et la simulation sont montrées dans la Fig. 3.30, avec leurs rapports, en fonction de E_T^{SC} , pour les différentes régions en η .

Les efficacités sont $> 85\%$ pour $E_T^{SC} > 10$ GeV pour les quatre régions en η présentées dans les figures. Dans la région de transition ($1.44 < |\eta| < 1.57$), elles sont inférieures à celles des régions voisines, différant de 10 à 20 % pour $10 < E_T^{SC} < 20$ GeV et de 1 à 10 % pour $E_T^{SC} > 20$ GeV. Les données sont compatibles avec les prédictions, avec des rapports des efficacités entre les données et la simulation consistants avec l'unité sur la quasi-totalité de la gamme en p_T et des régions en η . Les incertitudes montrées sur les figures correspondent à la somme quadratique des contributions statistiques et systématiques. Dominées par la composante systématique, elles s'élèvent à quelques pourcents pour $E_T^{SC} < 20$ GeV et diminuent jusqu'à $< 1\%$ lorsque E_T^{SC} augmente. L'incertitude dominante est liée au choix des fonctions d'ajustement. Par ailleurs, la contribution du bruit de fond étant importante dans l'estimation de l'efficacité de la reconstruction, des sélections supplémentaires sont imposées : l'énergie transverse manquante des événements est requise < 20 GeV, et le SC utilisé comme *probe* doit être isolé en requérant $\sum p_T/E_T^{SC} < 15\%$, la somme portant sur les traces dans le cône d'isolation provenant du vertex de l'événement. La définition de ces sélections supplémentaires est la deuxième source d'incertitude systématique pour cette mesure.

3.6.2 Efficacité de sélection

L'efficacité de sélection est mesurée pour des électrons reconstruits, par intervalle en p_T et η de l'électron, comme montré dans la Fig. 3.31 pour la sélection utilisée dans l'analyse du boson de Higgs en quatre leptons. Celle-ci étant optimisée pour $p_T > 7$ GeV, elle est mesurée dans cette même gamme.

La Fig. 3.32 représente cette même efficacité en fonction du nombre de vertex d'interaction reconstruits (comme mentionné précédemment, en moyenne 15 vertex d'interaction sont reconstruits, pour 21 événements d'empilement par événement pour les données enregistrées en 2012).

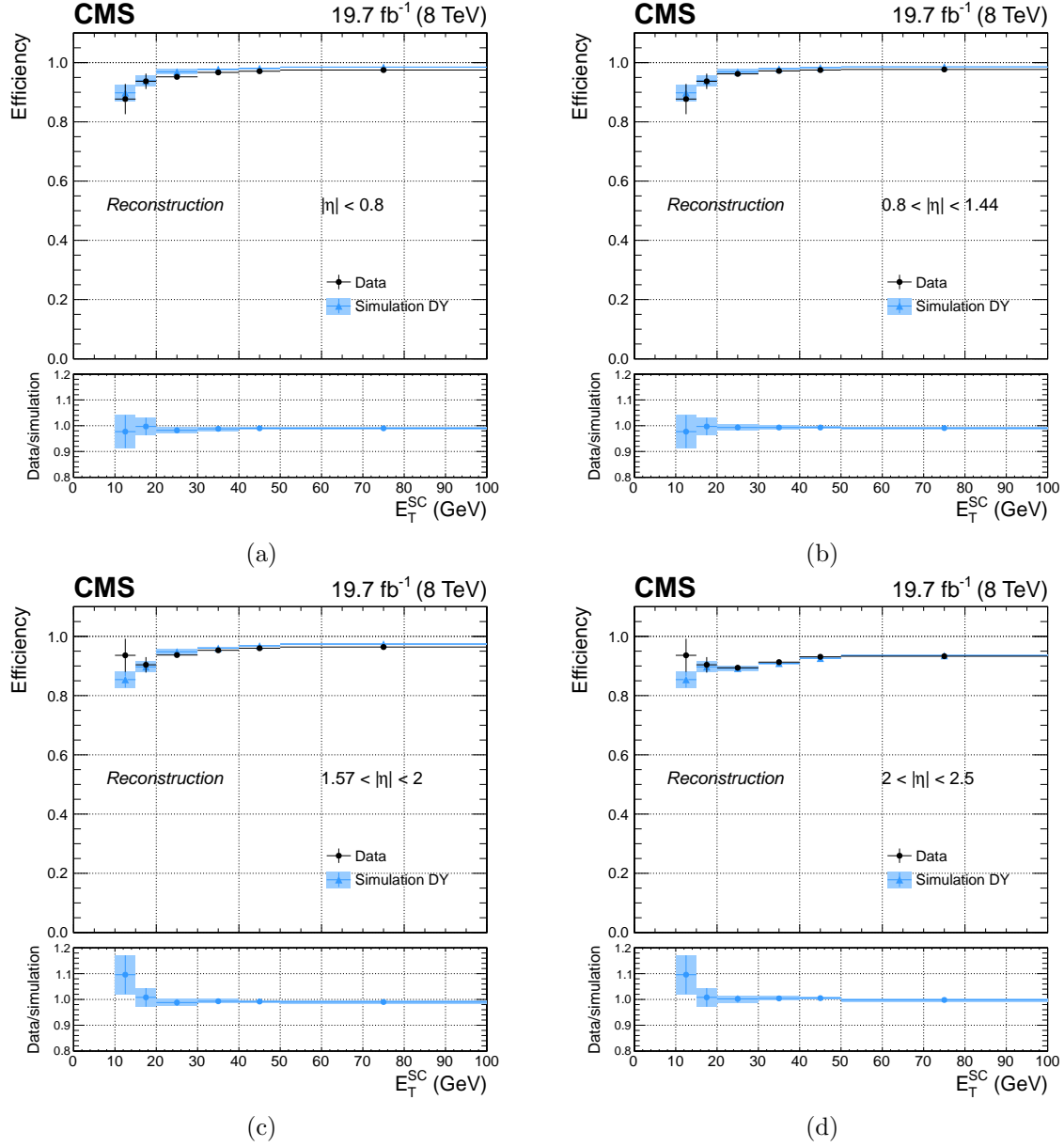


FIGURE 3.30 – Efficacité de reconstruction des électrons mesurée avec des événements $Z \rightarrow e^+e^-$ dans les données (points) et la simulation (triangles) en fonction du E_T^{SC} de l'électron pour (a) $|\eta| < 0.8$, (b) $0.8 < |\eta| < 1.44$, (c) $1.57 < |\eta| < 2$ et (d) $2 < |\eta| < 2.5$. Les figures sous les figures principales montrent les rapports entre les efficacités mesurées dans les données et la simulation. Les deux premiers intervalles, correspondant à $E_T^{\text{SC}} < 20$ GeV, sont identiques entre les deux régions du *barrel*, et entre les deux régions des *endcaps*, car ceux-ci ont été évalués de façon globale dans le *barrel* et dans les *endcaps*, à cause du faible nombre d'événements dans cette gamme en E_T .

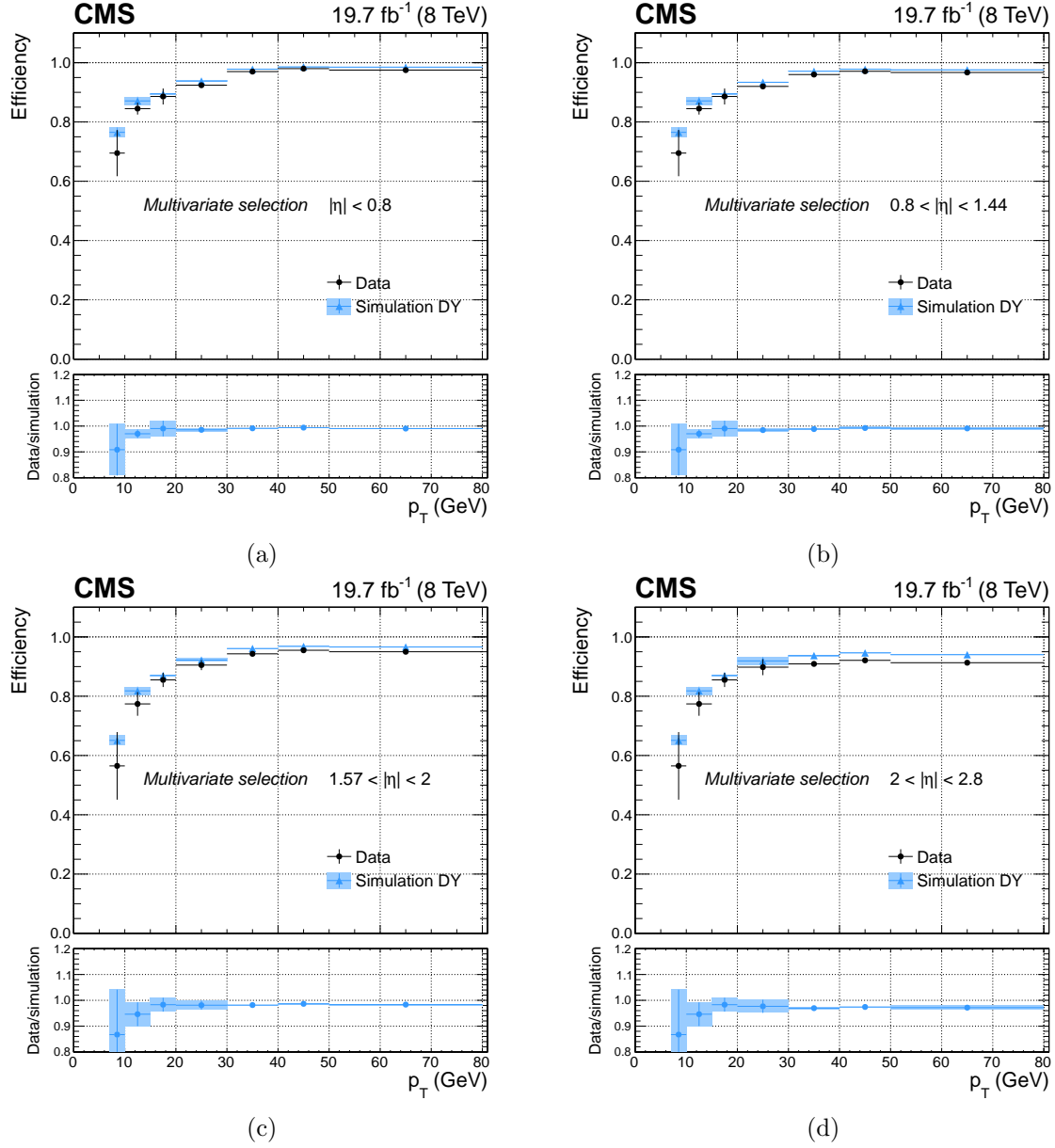


FIGURE 3.31 – Efficacité de la sélection des électrons utilisée dans l’analyse du boson de Higgs en quatre leptons, mesurée avec des événements $Z \rightarrow e^+e^-$ dans les données (points) et la simulation (triangles) en fonction du p_T de l’électron pour (a) $|\eta| < 0.8$, (b) $0.8 < |\eta| < 1.44$, (c) $1.57 < |\eta| < 2$ et (d) $2 < |\eta| < 2.5$. Les figures sous les figures principales montrent les rapports entre les efficacités mesurées dans les données et la simulation. Les trois premiers intervalles, correspondant à $p_T < 20$ GeV, sont identiques entre les deux régions du *barrel*, et entre les deux régions des *endcaps*, car ceux-ci ont été évalués de façon globale dans le *barrel* et dans les *endcaps*, à cause du faible nombre d’événements dans cette gamme en p_T .

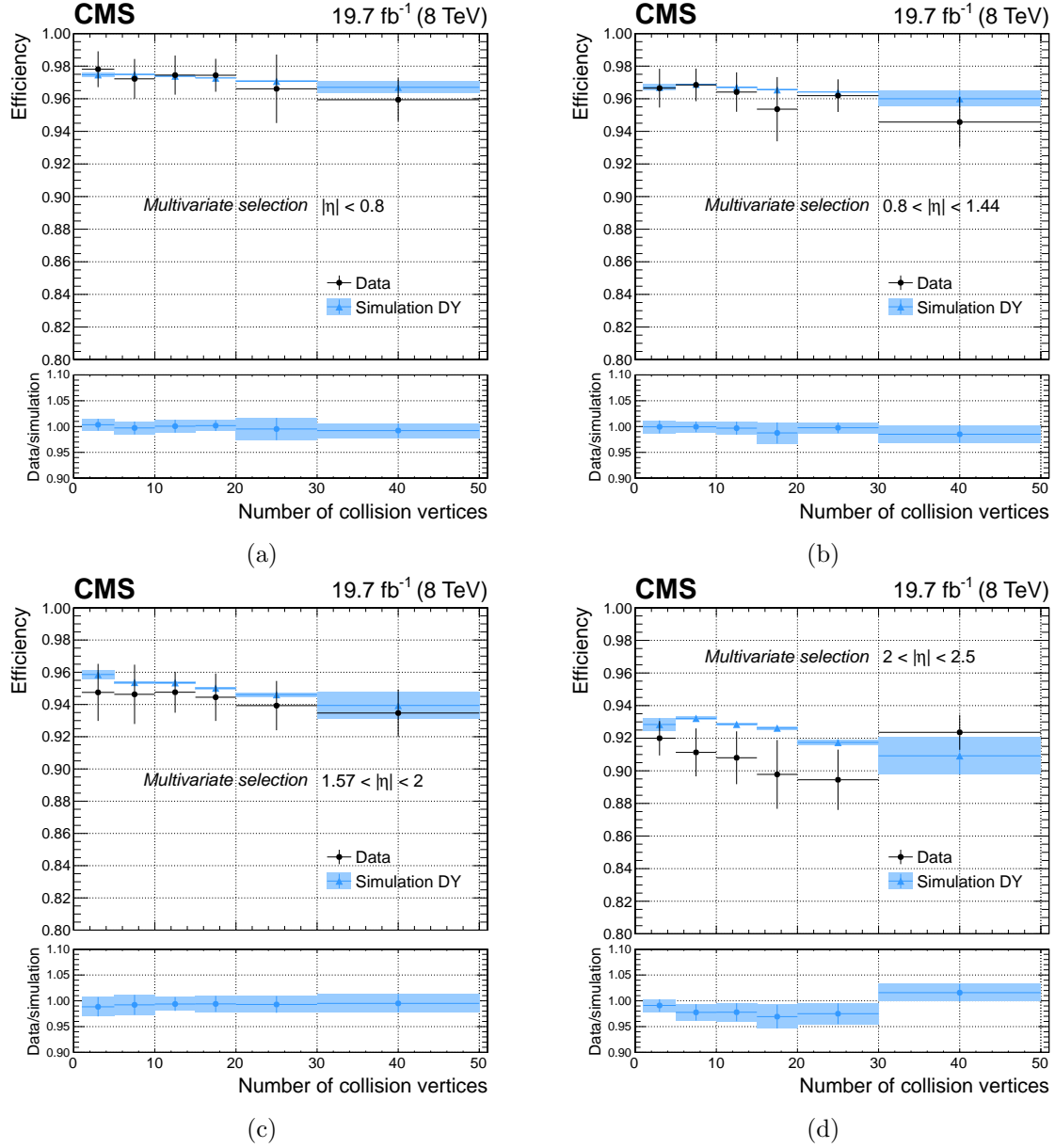


FIGURE 3.32 – Efficacité de la sélection des électrons utilisée dans l’analyse du boson de Higgs en quatre leptons, mesurée avec des événements $Z \rightarrow e^+e^-$ dans les données (points) et la simulation (triangles) en fonction du nombre de vertex d’interaction reconstruits pour (a) $|\eta| < 0.8$, (b) $0.8 < |\eta| < 1.44$, (c) $1.57 < |\eta| < 2$ et (d) $2 < |\eta| < 2.5$. Les figures sous les figures principales montrent les rapports entre les efficacités mesurées dans les données et la simulation.

De façon générale, les données et la simulation sont en bon accord. Les rapports entre les données et la simulation sont compatibles avec l'unité, hormis pour les régions à bas p_T ($7 < p_T < 15 \text{ GeV}$), où ils peuvent diminuer jusqu'à 0.85–0.90. Les efficacités sont quasi-uniformes en fonction du nombre de vertex reconstruits, diminuant seulement de $\lesssim 2\%$ en 50 vertex reconstruits, ce qui montre que la sélection n'est pas dépendante de l'empilement (ceci étant d'autant plus vrai que la sélection est relâchée). Néanmoins, une perte d'efficacité en fonction du nombre de vertex d'interactions légèrement plus importante, ainsi qu'un désaccord entre les données et la simulation, sont visibles dans la région la plus à l'avant, la plus sensible à l'empilement. Les incertitudes montrées sur les figures correspondent à la somme quadratique des contributions statistiques et systématiques et sont dominées par les incertitudes systématiques. Elles s'élèvent à quelques pourcents pour $p_T < 20 \text{ GeV}$, et diminuent jusqu'à 2% lorsque p_T augmente, à l'exception de la région de transition entre le *barrel* et les *endcaps* qui n'est pas illustrée dans ces figures. De même que pour l'efficacité de reconstruction, l'incertitude dominante provient du choix des fonctions d'ajustement.

3.6.3 Probabilité de fausse identification

A chaque sélection correspond une efficacité et une probabilité de fausse identification, définie comme la fraction des candidats de bruits de fond reconstruits en électrons passant le critère de sélection considéré. La probabilité de fausse identification d'une sélection donnée est l'ingrédient de base pour la mesure des bruits de fond réductibles dans l'analyse correspondante.

Pour la mesure de la probabilité de fausse identification de la sélection des électrons utilisée dans l'analyse du boson de Higgs en quatre leptons (telle que décrite dans la Table 3.5), l'échantillon de données contenant des événements Z et un jet reconstruit comme un électron, décrit dans la Section 3.5.1, est utilisé. Les résultats sont montrés dans la Fig. 3.33 en fonction du p_T du candidat électron. La probabilité de fausse identification augmente en fonction du p_T du candidat électron et vaut entre 1.5 to 10.5 %. Les incertitudes indiquées sur la figure correspondent uniquement à la contribution statistique.

Une source significative d'incertitude systématique est reliée à la composition de l'échantillon utilisé pour la mesure, et en particulier à la contamination de celui-ci par des électrons véritables issus de la production associée de W et Z, ou d'événements $t\bar{t}$. Cet effet augmente avec le p_T du candidat mais est largement réduit par la sélection sur l'énergie transverse manquante. En conséquence, l'incertitude provenant de la différence entre l'échantillon utilisé pour extraire la mesure et celui pour lequel celle-ci est appliquée devient prépondérante. Son évaluation et sa prise en compte sont décrites dans le chapitre suivant.

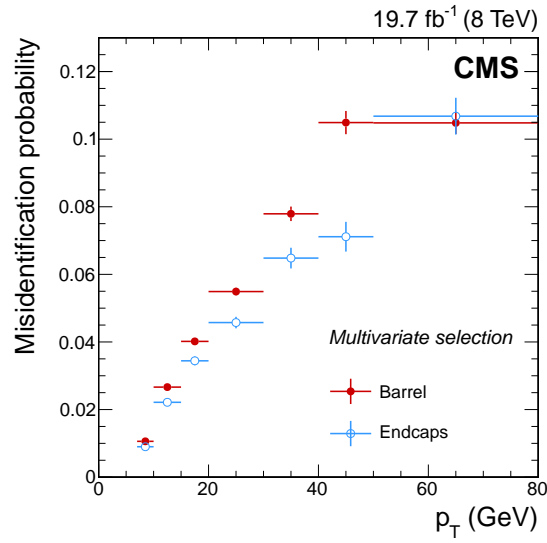


FIGURE 3.33 – Probabilité de fausse identification mesurée dans les données, en fonction du p_T du candidat électron dans le *barrel* (points rouges) et les *endcaps* (points bleus) pour la sélection des électrons utilisée dans l’analyse du boson de Higgs en quatre leptons.

Ce troisième chapitre, fondamental dans ce document et également concernant mon travail, a montré la description des algorithmes de reconstruction et de sélection des électrons et leurs performances. À présent, nous sommes donc pourvus de la totalité des ingrédients pour l'analyse du boson de Higgs dans son canal de désintégration en quatre leptons et en particulier la mesure de sa masse, objet du prochain chapitre.

Chapitre 4

Mesurer

Sommaire

4.1	Considérations générales	100
4.1.1	Stratégie	100
4.1.2	Signature et bruits de fond	100
4.1.3	Échantillons de données et de simulation	101
4.2	Sélection des événements	102
4.2.1	Leptons	103
4.2.2	Radiations dans l'état final	103
4.2.3	Sélections cinématiques	104
4.3	Estimation des bruits de fond	105
4.3.1	Bruit de fond irréductible ZZ	105
4.3.2	Bruit de fond réductible	106
4.3.2.1	Stratégie	106
4.3.2.2	Probabilités de fausse identification	107
4.3.2.3	Méthode OS	107
4.3.2.4	Méthode SS	108
4.3.2.5	Résultat	108
4.4	Discriminant cinématique	108
4.5	Petit historique de la découverte	112
4.6	Mesure de la masse	117
4.6.1	Modélisation de $m_{4\ell}$	117
4.6.2	Incertitudes sur $m_{4\ell}$ par événement	118
4.6.3	Discriminant cinématique	120
4.6.4	Incertitudes systématiques	120
4.6.5	Résultats	120
4.6.5.1	Mesure de m_H avec le canal en quatre leptons dans CMS	120
4.6.5.2	Combinaison ultime	122

À ce stade, la quasi-totalité des événements produits dans les collisions et contenant des bosons de Higgs en quatre leptons sont enregistrés, ainsi que beaucoup d'autres qui y ressemblent. Les outils nécessaires à la reconstruction, l'identification et l'isolation des leptons ont été développés et optimisés. Le premier traitement de ces événements est donc de s'assurer qu'ils ont bien la signature voulue à l'aide d'une sélection précise et adaptée, avant d'évaluer les bruits de fond restants et les erreurs systématiques. Enfin, lorsque la présence de ce boson de Higgs en quatre leptons est confirmée, la poignée de ces précieux événements est exploitée pour en déduire les meilleures mesures possibles de ses propriétés, et en particulier sa masse. La majorité des résultats exposés dans ce chapitre sont issus de la publication finale pour le canal en quatre leptons, dite “legacy” sur le sujet [34], et d'un des documents internes à CMS associés comprenant tous les détails [75]. Certaines informations complémentaires peuvent également être trouvées dans le manuscrit de thèse de Mykhailo Dalchenko [76]. Les résultats finaux concernant les mesures des propriétés du boson de Higgs pour la combinaison des tous les canaux dans CMS sont présentés dans la Ref. [77].

4.1 Considérations générales

4.1.1 Stratégie

La stratégie, les outils et ingrédients nécessaires à l'analyse du boson de Higgs en quatre leptons ont été développés et optimisés dans un premier temps pour la préparation du Physics TDR [78, 79] puis sans cesse améliorés jusqu'à la découverte en 2012 [80, 81]. Ils reposent sur une solide reconstruction des leptons, une connaissance et une maîtrise approfondie de l'échelle en impulsion des leptons, une efficacité de sélection maximale et un contrôle précis des bruits de fond. La ligne de conduite consiste en la recherche d'une résonance étroite au-dessus d'un bruit de fond continu dans la distribution de la masse invariante des quatre leptons. Les premiers pas avec les données ont été consacrés à la mise en service (*commissioning*) des objets (en particulier pour les électrons, voir [82–84], et [85–90] pour les documents internes¹), puis à l'analyse des premiers événements avec la signature recherchée, avant l'analyse finale. Le concept de *blinding*, qui consiste à optimiser les outils et la stratégie sur les simulations et de n'observer les données qu'en dehors de la zone de signal, a été adopté très en amont lors de la préparation de l'analyse.

4.1.2 Signature et bruits de fond

Les bosons de Higgs se désintégrant en quatre leptons laissent dans CMS une signature très claire : deux paires de leptons de même saveur et de charges opposées, primaires et isolés, et formant deux bosons Z dont l'un ou les deux peuvent se trouver hors couche de masse. Un exemple de cette signature peut être observé dans la Fig. 4.1, représentation graphique d'un événement candidat de boson de Higgs en deux électrons et deux muons, enregistré en mai 2012. Seuls les électrons et muons sont considérés, les leptons taus possédant une signature

1. La liste d'auteurs des documents internes est ordonnée par ordre alphabétique.

beaucoup plus complexe et une résolution moins précise.

Aussi significative que soit cette signature, elle peut néanmoins incomber à d'autres événements non issus de bosons de Higgs, soit de façon native (bruit de fond irréductible), soit par l'incursion de fausses identifications d'autres particules issues de l'activité hadronique inhérente aux collisions de protons (bruit de fond réductible). Seuls les bruits de fond du modèle standard sont considérés. Le bruit de fond dominant est irréductible, il provient de la production de ZZ ou $Z\gamma^*$ par l'annihilation d'un quark et d'un antiquark ($q\bar{q} \rightarrow ZZ$) ou la fusion de gluons ($gg \rightarrow ZZ$), suivie de la désintégration leptonique des Z ou γ^* , et est nommé ZZ par la suite. Les contributions suivantes (réductibles) proviennent des productions de Z +jets (jets légers, $b\bar{b}$, ou $c\bar{c}$), ou de $t\bar{t}$, suivies de la désintégration leptonique du Z ou des W issus des tops, avec la présence de leptons secondaires issus des désintégrations dans les jets de saveurs lourdes (b ou c), ou de leptons faussement identifiés provenant des mésons légers, et en particulier pour les électrons de la superposition de photons avec des hadrons chargés. D'autres bruits de fond proviennent de la production de dibosons ($Z\gamma$, WW , WZ) accompagnés de jets, suivie de la désintégration leptonique des Z et W , avec des leptons faussement identifiés dans les jets.

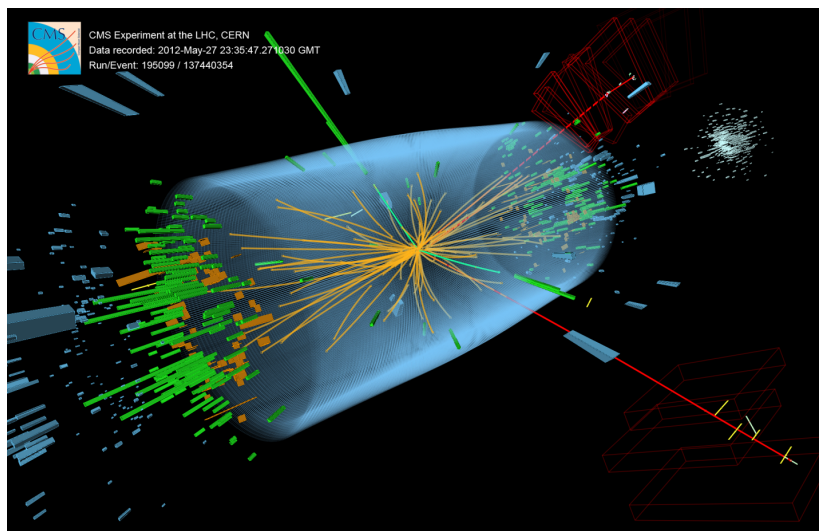


FIGURE 4.1 – Représentation d'un événement candidat de boson de Higgs en deux électrons et deux muons, enregistré par CMS en mai 2012.

4.1.3 Échantillons de données et de simulation

Les données exploitées représentent une luminosité intégrée égale à 24.8 fb^{-1} , correspondant à 5.1 fb^{-1} pour une énergie dans le centre de masse égale à 7 TeV (en 2011) et 19.7 fb^{-1} pour une énergie égale à 8 TeV (en 2012). Les échantillons de simulation pour le signal et

les bruits de fond sont générés à l’aide de logiciels basés sur les calculs théoriques les plus récents et les plus complets, puis interfacés avec PYTHIA6.424 pour la description des radiations dans les états initial et final, de l’hadronisation des partons, et de l’événement sous-jacent (avec les paramètres PYTHIA tune Z2 et Z2* pour 7 et 8 TeV, respectivement). Les événements additionnels dus à l’empilement d’interactions proton-proton lors d’un même croisement de faisceaux sont également générés et ajoutés (en moyenne 9 et 21 pour 7 et 8 TeV, respectivement). Les signatures des événements dans le détecteur sont ensuite engendrées grâce à la simulation complète et détaillée de CMS basée sur GEANT4, et enfin les mêmes déclenchement et reconstruction que pour les données leur sont appliquées. Les échantillons de simulation sont utilisés pour l’optimisation de la sélection et l’évaluation de l’acceptance et des incertitudes systématiques.

Les échantillons de signal produit par fusion de gluons (gg) ou fusion de bosons vecteurs (VBF) sont générés avec POWHEG [91–93] à l’ordre *next-to-leading order* (NLO) de QCD, les échantillons de signal avec production associée de W, Z ou $t\bar{t}$ sont générés avec PYTHIA. Les sections efficaces sont repondérées afin d’inclure les contributions venant de gg jusqu’à l’ordre *next-to-next-to-leading order* (NNLO) et *next-to-next-to-leading logarithm* (NNLL) [94–106], et celles venant de VBF jusqu’au NNLO [98, 107–111]. L’interférence entre le signal et le bruit de fond ZZ produits par gg est prise en compte [112]. La distribution en masse invariante du boson de Higgs est décrite avec une fonction Breit-Wigner (BW) pour $m_H < 400$ GeV. La désintégration du boson de Higgs est modélisée avec JHUGEN 3.1.8 [113–115] et inclut les effets d’interférence associés aux permutations des leptons identiques dans les canaux en quatre muons et en quatre électrons.

Les échantillons de bruit de fond irréductible ZZ sont générés par POWHEG [116] au NLO pour la production $q\bar{q} \rightarrow ZZ$, par GG2ZZ(LO) [117] pour $gg \rightarrow ZZ$, et leurs sections efficaces sont calculées par MCFM [118–120]. La contribution relative de $gg \rightarrow ZZ$ (LO) par rapport à $q\bar{q} \rightarrow ZZ$ (NLO) est $\approx 2\%$ pour une masse invariante des quatre leptons ($m_{4\ell}$) égale à 126 GeV et $\approx 6\%$ pour $m_{4\ell} = 1$ TeV. Les principaux échantillons de production de Z associés à d’autres particules, tels que $Zb\bar{b}$, $Zc\bar{c}$, $Z\gamma$ et Z+jets de quarks légers (dont l’ensemble est nommé Z+jets par la suite), sont générés par MADGRAPH [48], comprennent la production inclusive de Z avec jusqu’à quatre partons additionnels au niveau de l’élément de matrice, et leurs sections efficaces sont normalisées au calcul au NNLO. Les événements $t\bar{t}$ sont générés au NLO avec POWHEG. Les processus WW et WZ sont générés avec MADGRAPH et normalisés aux sections efficaces calculées au NLO.

4.2 Sélection des événements

La première sélection est constituée par le déclenchement, décrit en détail dans le chapitre 2, qui sélectionne les événements avec deux leptons ou plus avec une efficacité $> 99\%$ pour des événements de Higgs simulés avec une masse $m_H > 120$ GeV se désintégrant en

quatre leptons. Elle est suivie par la requête d’au moins un vertex d’interaction valide, c’est-à-dire éloigné du point central de CMS d’un maximum de 24 cm en z et de 2 cm dans le plan transverse, et à l’origine d’au moins quatre traces reconstruites. Les sélections suivantes concernent les leptons et les considérations cinématiques.

4.2.1 Leptons

La seconde étape concerne la sélection des leptons. Les électrons ont été traités en détails dans le chapitre 3. Ils sont sélectionnés avec $p_T^e > 7 \text{ GeV}$ et $|\eta^e| < 2.5$, et avec l’algorithme comprenant l’identification BDT pour les électrons *not-triggering*, l’isolation PF et la réjection des conversions exposées dans la Table 3.5. Comme décrit dans la Section 3.4.4, l’échelle en impulsion des électrons est connue avec une incertitude de 0.2 % dans le *barrel* et 0.3 % dans les *endcaps* (après l’application de toutes les corrections) et leur résolution effective est comprise entre 1.7 % pour des électrons bien mesurés avec un bremsstrahlung faible dans le *barrel* et 4.5 % pour des électrons avec un bremsstrahlung important ou mal mesurés dans les *endcaps*.

Concernant les muons, peu abordés dans ce mémoire, leur reconstruction combine les informations du trajectographe et des détecteurs à muons [121]. L’association entre les traces du trajectographe et celles reconstruites dans les chambres à muons est effectuée dans les deux sens, de l’intérieur vers l’extérieur de CMS et vice-versa. Certains muons sont également sélectionnés si une trace dans le trajectographe est associée à une collection partielle de segments de traces dans les chambres à muons, afin d’augmenter l’efficacité à bas p_T . Les muons sont sélectionnés avec $p_T^\mu > 5 \text{ GeV}$ et dans leur acceptation géométrique définie par $|\eta^\mu| < 2.4$, et identifiés à l’aide de critères de qualité des traces dans les deux sous-systèmes, et de compatibilité entre les traces et également avec des dépôts faibles d’énergie dans les calorimètres. Comme pour les électrons, l’échelle en impulsion et la résolution sont contrôlées à l’aide d’événements $J/\psi \rightarrow \mu^+\mu^-$, $\Upsilon \rightarrow \mu^+\mu^-$, et $Z \rightarrow \mu^+\mu^-$. L’échelle en impulsion des muons exploités dans cette analyse est connue avec une incertitude de 0.1 %, et leur résolution en p_T est comprise entre 1.3 % et 2.0 % dans le *barrel*, et jusqu’à 6 % dans les *endcaps*.

4.2.2 Radiations dans l’état final

Dans le contexte particulier de la recherche du boson de Higgs en quatre leptons, où chaque événement compte, il est important de reconstruire correctement un boson Z dont un lepton subit une radiation dans l’état final (FSR). Les photons de basse énergie sont reconstruits efficacement par l’algorithme PF. Les photons de FSR ont la caractéristique d’être produits majoritairement dans une direction colinéaire au lepton parent, et avec une distribution en p_T plus haute que les photons provenant de radiations dans l’état initial ou des événements d’empilement. Ainsi la sélection optimisée des photons FSR considère un seuil en p_T minimal de 2 GeV pour un photon très proche d’un lepton ($\Delta R < 0.07$), et un seuil

de 4 GeV adjoint d'un critère d'isolation PF, $\text{Iso}_{\text{PF}}/p_{\text{T}}^{\gamma} < 1$, Iso_{PF} étant défini comme dans la Section 3.5.2 avec $\Delta R = 0.3$, pour un photon un peu plus lointain ($0.07 < \Delta R < 0.5$). Un seul candidat FSR est associé à un Z. Dans le cas où plusieurs sont sélectionnés, le photon choisi est celui avec le plus haut p_{T}^{γ} si celui-ci est > 4 GeV, celui le plus proche d'un des leptons du Z dans le cas contraire. Enfin un candidat FSR n'est sélectionné que s'il permet à la masse d'un boson Z de s'approcher de sa masse nominale sans toutefois dépasser 100 GeV. L'impulsion d'un photon FSR sélectionné est ajoutée au lepton considéré et retiré des observables d'isolation. Les événements Higgs en quatre leptons simulés sont affectés par cette procédure à raison de 1.5, 4.6, et 9 % pour les canaux en quatre électrons, deux électrons et deux muons, et quatre muons, respectivement. Le gain en efficacité attendu est égal à 1, 2 et 3 %. Cette procédure affecte moins les électrons, les SC incluant par construction une fraction non négligeable des photons FSR.

4.2.3 Sélections cinématiques

La signature du boson de Higgs en quatre leptons conditionne la sélection première des événements : deux paires de leptons identifiés et isolés, de même saveur et de charges opposées. Les quatre leptons doivent provenir du même vertex d'interaction afin de limiter les événements de bruit de fond dont les leptons sont potentiellement secondaires. Afin que les p_{T} des leptons ne se situent pas dans une phase de changement trop rapide de l'efficacité des *triggers*, un des leptons doit vérifier $p_{\text{T}} > 20$ GeV, et un autre $p_{\text{T}} > 10$ GeV. La première paire sélectionnée, nommée Z1, est celle avec la masse invariante la plus proche de la masse nominale du Z, vérifiant $40 < m_{\text{Z1}} < 120$ GeV, en tenant compte des photons FSR. La deuxième paire, nommée Z2, vérifie $12 < m_{\text{Z2}} < 120$ GeV. Dans le cas où plusieurs candidats Z2 sont possibles, le choix se porte sur celui avec la plus grande somme scalaire des p_{T} de ses leptons. Le seuil minimal pour m_{Z2} est maintenu le plus bas possible afin de conserver les bosons de Higgs éventuels avec $m_{\text{H}} < 180$ GeV, pour lesquels au moins un boson Z est hors couche de masse. Parmi les quatre leptons élus, toutes les masses invariantes des paires de charges opposées sont requises > 4 GeV afin d'éviter la sélection de bruits de fond constitués de résonances à basse masse ou de désintégrations hadroniques dans les jets. L'espace de phase pour la recherche du boson de Higgs est limité à $m_{4\ell} > 100$ GeV.

Cette sélection, aussi simple semble-t-elle, a été approfondie dans ses moindres détails dans le souci constant d'une efficacité et d'une pureté maximales. Notamment, la sélection en $p_{\text{T}} > 20, 10$ GeV appliquée parmi les quatre leptons formant le candidat Higgs, a été améliorée par rapport à la sélection initiale qui était appliquée sur les leptons du Z1, permettant ainsi un gain relatif de 1 – 3 % d'efficacité de sélection des événements. De même, l'étude du choix de l'ensemble des quatre leptons conclut que celui-ci est correct dans > 99.9 % des événements simulés de boson de Higgs dont les leptons sont reconstruits (dans le cas de la production par gg).

Les efficacités de sélection sont illustrées dans la Fig. 4.2 en fonction de m_{H} , pour des bosons de Higgs simulés avec les quatre leptons dans l'acceptance du détecteur, et produits

par fusion de gluons avec $\sqrt{s} = 8 \text{ TeV}$ (les autres modes présentent des efficacités similaires). Celles-ci augmentent significativement jusqu'à $m_H = 2m_Z$ puis atteignent un plateau. La présence de quatre leptons implique une efficacité pour le canal en quatre muons, pour lesquels l'efficacité de reconstruction et d'identification est meilleure que celle des électrons, bien supérieure à celle du canal en quatre électrons. Les efficacités pour $m_H = 126 \text{ GeV}$ valent 30, 43, et 62 % pour les canaux $4e$, $2e2\mu$ et 4μ , respectivement. Elles valent 58, 71, et 87 % pour $m_H = 200 \text{ GeV}$.

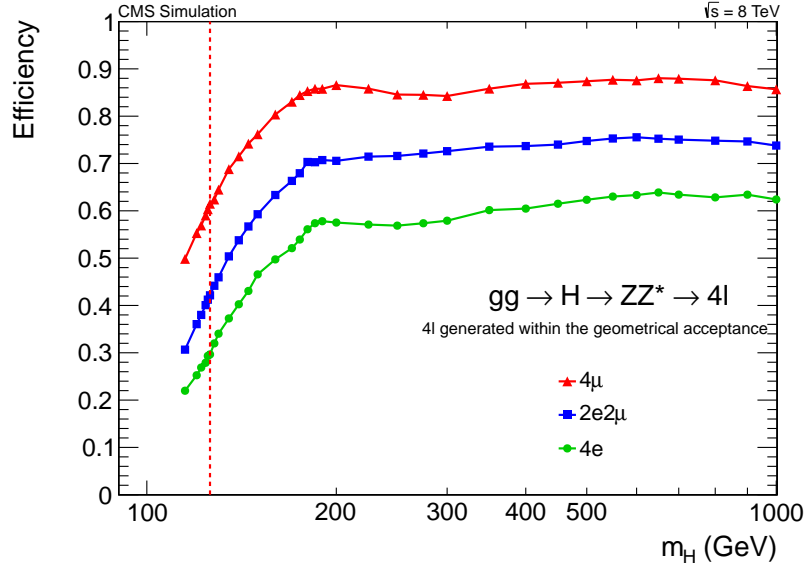


FIGURE 4.2 – Efficacités de sélection des bosons de Higgs produits dans l'acceptance géométrique, par fusion de gluons avec une énergie dans le centre de masse égale à 8 TeV, en fonction de m_H pour les trois canaux de désintégration en quatre leptons (courbe verte : $4e$, courbe bleue : $2e2\mu$, courbe rouge : 4μ).

4.3 Estimation des bruits de fond

4.3.1 Bruit de fond irréductible ZZ

La production d'une paire de bosons Z représente le bruit de fond le plus important. Il contribue, subséquentement à la totalité de la sélection et respectivement pour les canaux $4e$, $2e2\mu$, et 4μ , à hauteur de 91, 94, et 97 % du bruit de fond total dans la région $100 < m_{4\ell} < 1000 \text{ GeV}$, et à hauteur de 58, 71, et 86 % dans la région $121.5 < m_{4\ell} < 130.5 \text{ GeV}$. La quantité attendue de tels événements ainsi que leurs distributions cinématiques sont estimées par la simulation.

L'incertitude systématique majeure pour ce bruit de fond réside dans la normalisation. L'incertitude théorique est évaluée en fonction de $m_{4\ell}$ [122–126] : elle s'accroît avec $m_{4\ell}$, de 4 à 14 % pour $q\bar{q} \rightarrow ZZ$ (et vaut 4.5 % pour une analyse avec $m_H = 126$ GeV) et de 25 à 50 % pour $gg \rightarrow ZZ$ (et vaut 25.1 % pour $m_H = 126$ GeV). Les bruits de fond provenant d'une interaction double de partons, chacune à l'origine de la production d'un boson Z, ont été évalués et leur contribution est négligeable.

Les incertitudes expérimentales proviennent de la précision sur la mesure de la luminosité d'une part, et de la différence d'efficacité de sélection entre les données et la simulation d'autre part. La luminosité est connue avec une précision de 2.2 % [127] en 2011 et 2.6 % en 2012 [45]. Les différences entre les données et la simulation pour la sélection concernent le déclenchement et la reconstruction, l'identification, et l'isolation des leptons. En ce qui concerne le déclenchement, aucune correction n'est appliquée à la simulation et, comme discuté au chapitre 2, l'incertitude sur le nombre d'événements déclenchés est égale à 1.5 %. Pour la reconstruction, l'identification, et l'isolation, les mesures individuelles par lepton obtenues par la méthode de T&P sont exploitées et propagées dans la simulation, en particulier les rapports d'efficacité entre les données et la simulation (*ScaleFactor*). Pour chacun des leptons un facteur de correction, résultat d'un tirage aléatoire avec une distribution gaussienne centrée sur le *ScaleFactor* (mesuré par intervalle de p_T et η) et de largeur l'erreur associée, est appliqué comme un poids multiplicatif dans l'événement. Ces pseudo-expériences simulées sont itérées 500 fois, l'erreur systématique sur le nombre d'événements attendu est alors le RMS de la distribution obtenue. Celle-ci diminue en fonction de $m_{4\ell}$ et vaut de 5.5 à 11 % pour le canal $4e$ (10 % pour $m_H = 126$ GeV), et de 2.9 à 4.3 % pour le canal 4μ (4.3 % pour $m_H = 126$ GeV).

4.3.2 Bruit de fond réductible

4.3.2.1 Stratégie

Le bruit de fond réductible est estimé à l'aide des données. En effet, les effets de fausse identification des leptons étant rares et complexes, une simulation correcte nécessiterait une statistique incommensurable et une connaissance théorique et expérimentale extrêmement précise de chaque source de fausse identification. La stratégie générale est d'utiliser les données pour mesurer les probabilités de fausse identification d'une part, et d'autre part de les appliquer dans des régions de contrôle suffisamment proches de la zone de signal de l'analyse tout en étant dépourvues. Toute la subtilité et la complexité de cette estimation réside dans le choix des régions de contrôle et de mesure des probabilités de fausse identification, dans la connaissance de leur composition en bruits de fond et de leur contamination en signal. Deux méthodes différentes et indépendantes sont exploitées.

Les régions de contrôle sont définies par des événements contenant une paire de leptons formant un candidat Z1 tel que dans la sélection du signal, et deux leptons additionnels,

de charges opposées (méthode OS) ou identiques (méthode SS) satisfaisant des sélections relâchées par rapport à celle du signal. La masse invariante de ces leptons additionnels est requise $> 12 \text{ GeV}$ afin d'être compatible avec la sélection du signal.

4.3.2.2 Probabilités de fausse identification

Les probabilités de fausse identification sont évaluées avec des événements contenant une paire de leptons formant un candidat Z1, et un unique lepton additionnel, vérifiant des sélections relâchées et provenant majoritairement du bruit de fond. Elles consistent en la fraction des leptons additionnels qui vérifient les sélections de l'analyse en fonction de leur p_T et η . Un exemple d'une telle mesure pour les électrons est décrite dans la Section 3.6.3. Les résultats dépendent fortement des sélections des événements et du lepton additionnel, car celles-ci définissent la composition en bruit de fond considérée. Une énergie transverse manquante $< 25 \text{ GeV}$ est requise pour éliminer les contributions des productions WZ et $t\bar{t}$. La masse invariante entre le lepton additionnel et le lepton de charge opposée du Z1, s'ils sont de même saveur, doit être $> 4 \text{ GeV}$ pour rejeter les résonances à basse masse, telles que le J/ψ . La masse invariante du Z1 ne doit pas s'éloigner de la masse nominale du Z de plus de 10 GeV pour la méthode OS, et doit être comprise entre 40 et 120 GeV pour la méthode SS. De ce fait la composition de cet échantillon est différente pour les deux méthodes, en particulier en ce qui concerne la présence de FSR dans les événements. Les probabilités de fausse identification ainsi obtenues sont comprises de 1 à 15% pour les électrons, et de 5 à 10% pour les muons.

4.3.2.3 Méthode OS

Pour cette méthode la région de contrôle est définie par des événements contenant un candidat Z1 et deux leptons additionnels de même saveur et de charges opposées. Une première catégorie est définie pour les événements dont les deux leptons additionnels ne vérifient pas les sélections de l'analyse et permet d'estimer les bruits de fond ayant intrinsèquement deux leptons primaires (Z+jets, $t\bar{t}$) en pondérant les événements à l'aide des probabilités de fausse identification de chacun des leptons additionnels. La deuxième catégorie est définie pour les événements pour lesquels un seul lepton additionnel vérifie les sélections de l'analyse et permet d'estimer les bruits de fond ayant intrinsèquement trois leptons réels (WZ+jets), ou presque ($Z\gamma$ +jets avec conversion du photon) en pondérant les événements à l'aide des probabilités de fausse identification du lepton additionnel ne passant pas les sélections de l'analyse. L'estimation totale du bruit de fond réductible somme les deux contributions. Des corrections sont apportées, à cause de la présence d'une part d'événements ZZ dans la deuxième catégorie, lorsqu'un lepton ne passe pas les sélections de l'analyse, et d'autre part d'événements avec deux leptons primaires mais entrant dans la deuxième catégorie. L'utilisation des deux catégories permet de compenser l'absence de photons convertis dans l'estimation des probabilités de fausse identification.

4.3.2.4 Méthode SS

Pour cette méthode la région de contrôle est définie par des événements contenant un candidat Z1 et deux leptons additionnels de même saveur et de même charge. La sélection sur la masse invariante du candidat Z1 dans l'échantillon de la mesure de probabilité de fausse identification étant beaucoup plus relâchée que pour la méthode OS, cette méthode prend en compte les événements avec FSR et les événements $Z + \gamma$. De plus, elle présente une sélection plus proche que celle de l'analyse. Cependant, certains photons de FSR qui convertissent augmentent sensiblement les probabilités de fausse identification, et la fréquence des FSR est différente entre l'échantillon de mesure des probabilités de fausse identification et l'échantillon de contrôle, apportant ainsi un biais dans la mesure du bruit de fond réductible. Cet effet est corrigé par un facteur dérivé de la dépendance linéaire observée de la probabilité de fausse identification en fonction de la fraction d'électrons avec un *hit* interne manquant (cette observable directement liée à la conversion de photon est expliquée en détails dans la Section 3.3.5). Les événements sont pondérés à l'aide des probabilités de fausse identification ainsi corrigées pour chacun des leptons additionnels. Le bruit de fond réductible total est alors estimé en multipliant le résultat par le rapport du nombre d'événements avec deux leptons additionnels de charges opposées sur le nombre d'événements avec deux leptons additionnels de même charge, obtenu par la simulation.

4.3.2.5 Résultat

Les nombres de bruits de fond prédits par les deux méthodes sont compatibles en tenant compte de leurs erreurs statistiques, en conséquence ils sont combinés pour obtenir le résultat final. La forme de la distribution en $m_{4\ell}$ est estimée par un ajustement empirique des événements de chaque catégorie de la méthode OS, avec comme incertitude systématique l'enveloppe couvrant des ajustements alternatifs. L'incertitude systématique totale assignée au bruit de fond réductible comprend l'incertitude sur la distribution et la différence de composition entre les échantillons de mesure des probabilités de fausse identification et les régions de contrôle ; elle est estimée à 20, 25, et 40 % pour les canaux $4e$, $2e2\mu$, et 4μ , respectivement. Les deux méthodes ont été validées en utilisant des événements sélectionnés de la même façon que pour l'analyse, à l'exception du candidat Z2 formé par des leptons de saveurs opposées ou de même charge. La prédiction finale du bruit de fond réductible est montrée dans la Fig. 4.3. Celui-ci contribue, respectivement pour les canaux $4e$, $2e2\mu$, et 4μ , à hauteur de 9, 6, et 3 % du bruit de fond total dans la région $100 < m_{4\ell} < 1000$ GeV, et à hauteur de 42, 28 et 14 % dans la région $121.5 < m_{4\ell} < 130.5$ GeV.

4.4 Discriminant cinématique

L'un des avantages de l'analyse du boson de Higgs en quatre leptons est la détection de la totalité des produits de la désintégration du boson de Higgs, et ainsi la possibilité d'utiliser les observables cinématiques, notamment dans le but d'améliorer l'identification du signal par rapport au bruit de fond. De la meilleure sensibilité au signal découle la diminution

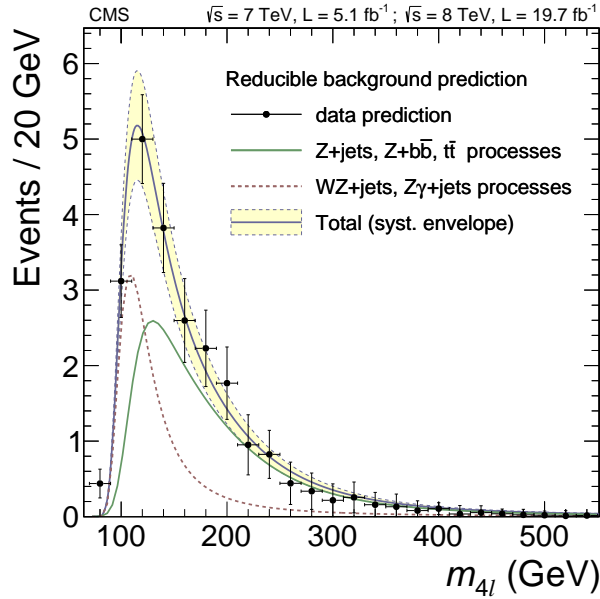


FIGURE 4.3 – Prédiction du bruit de fond réductible pour l’ensemble des 3 canaux de l’analyse du boson de Higgs en quatre leptons (points noirs) et son ajustement (courbe continue bleue) en fonction de $m_{4\ell}$. L’incertitude systématique totale, ainsi que les ajustements correspondant aux composantes avec deux (courbe continue verte) et trois (courbe en pointillé marron) leptons primaires intrinsèques sont également montrés.

des incertitudes statistiques associées aux mesures des propriétés du boson de Higgs. L’ensemble des caractéristiques cinématiques d’un événement de boson de Higgs se désintégrant en quatre leptons peuvent être décrites dans le centre de masse du boson de Higgs par cinq angles (deux définissant la direction des bosons Z, trois définissant les directions des leptons) et les masses invariantes des bosons Z. L’impulsion transverse et la position en η du boson de Higgs permettent d’obtenir les informations dans le référentiel du laboratoire mais ne sont pas utilisées dans la construction de l’observable discriminante, l’un étant sensible au mode de production et l’autre n’apportant pas de pouvoir discriminant supplémentaire significatif. Les distributions de masse invariante des candidats Z sont illustrées dans la Fig. 4.4 pour m_{Z1} (a), m_{Z2} (b) et m_{Z2} en fonction de m_{Z1} (c) après la sélection d’analyse pour le signal avec $m_H = 126$ GeV, les bruits de fonds et les données dans la gamme $121.5 < m_{4\ell} < 130.5$ GeV.

Le discriminant cinématique est construit à partir des probabilités d’un événement donné d’être du signal ($P_{\text{sig}}^{\text{kin}}$) ou du bruit de fond ($P_{\text{bkg}}^{\text{kin}}$), vis-à-vis de ses caractéristiques cinématiques pour une masse $m_{4\ell}$ donnée. Pour le signal et le bruit de fond ZZ, les distributions de probabilité cinématiques (angles et masses invariantes des bosons Z) sont calculées à l’aide des éléments de matrice au LO. Les éléments de matrice pour le signal sont calculés avec l’hypothèse $m_{4\ell} = m_H$. Pour le bruit de fond ZZ, seule la production dominante $q\bar{q} \rightarrow ZZ$ est

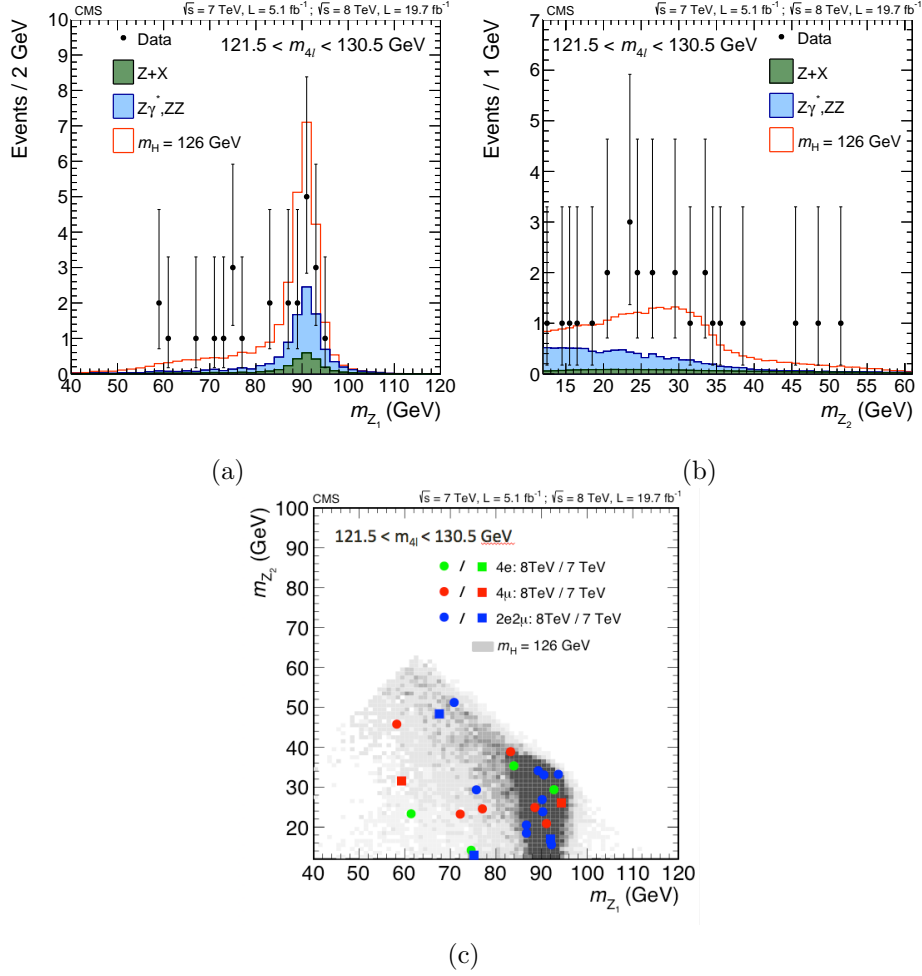


FIGURE 4.4 – Distributions des masses invariantes (a) de Z1, (b) de Z2, et (c) de Z2 en fonction de Z1 pour le signal avec $m_H = 126$ GeV, les bruits de fond ((a) et (b)), et les données dans la gamme $121.5 < m_{4\ell} < 130.5$ GeV après les sélections d'analyse.

considérée pour la paramétrisation des probabilités. Pour le bruit de fond réductible, les distributions de probabilité sont modélisées en utilisant les données dans les échantillons de contrôle précédemment décrits.

L'observable discriminante est alors définie par $D_{\text{kin}} = P_{\text{sig}}^{\text{kin}} / (P_{\text{sig}}^{\text{kin}} + P_{\text{bkg}}^{\text{kin}})$. Celle-ci n'inclut pas $m_{4\ell}$ qui est utilisée en observable discriminante supplémentaire. Les probabilités sont normalisées par des facteurs constants pour une valeur donnée de $m_{4\ell}$, de telle façon que la probabilité d'obtenir $D_{\text{kin}} > 0.5$ dans le cas du signal soit égale à la probabilité d'obtenir $D_{\text{kin}} < 0.5$ dans le cas du bruit de fond. Cette analyse utilise le logiciel MELA (*Matrix-Element Likelihood Approach*) [114, 115] avec le calcul des éléments de matrice obtenu avec JHUGEN pour le signal et avec MCFM pour le bruit de fond $q\bar{q} \rightarrow ZZ$. Les effets d'accep-

tance du détecteur deviennent négligeables dans les rapports de probabilités du discriminant, grâce notamment aux excellentes résolutions obtenues pour les leptons.

La Fig. 4.5 montre les distributions du discriminant cinématique ainsi construit, pour le signal avec $m_H = 126 \text{ GeV}$, les bruits de fond et les données dans la gamme $121.5 < m_{4\ell} < 130.5 \text{ GeV}$ (a), et en fonction de $m_{4\ell}$ pour les données et les bruits de fond (b), ainsi que pour les données et la somme des bruits de fond et du signal (c) pour la gamme $114 < m_{4\ell} < 180 \text{ GeV}$. Le pouvoir de discrimination ainsi que la tendance des données sont significativement visibles. Ce discriminant est utilisé pour toutes les analyses concernant le boson de Higgs en quatre leptons, et en particulier pour les exclusions, les probabilités de fluctuation de bruit de fond, et la masse.

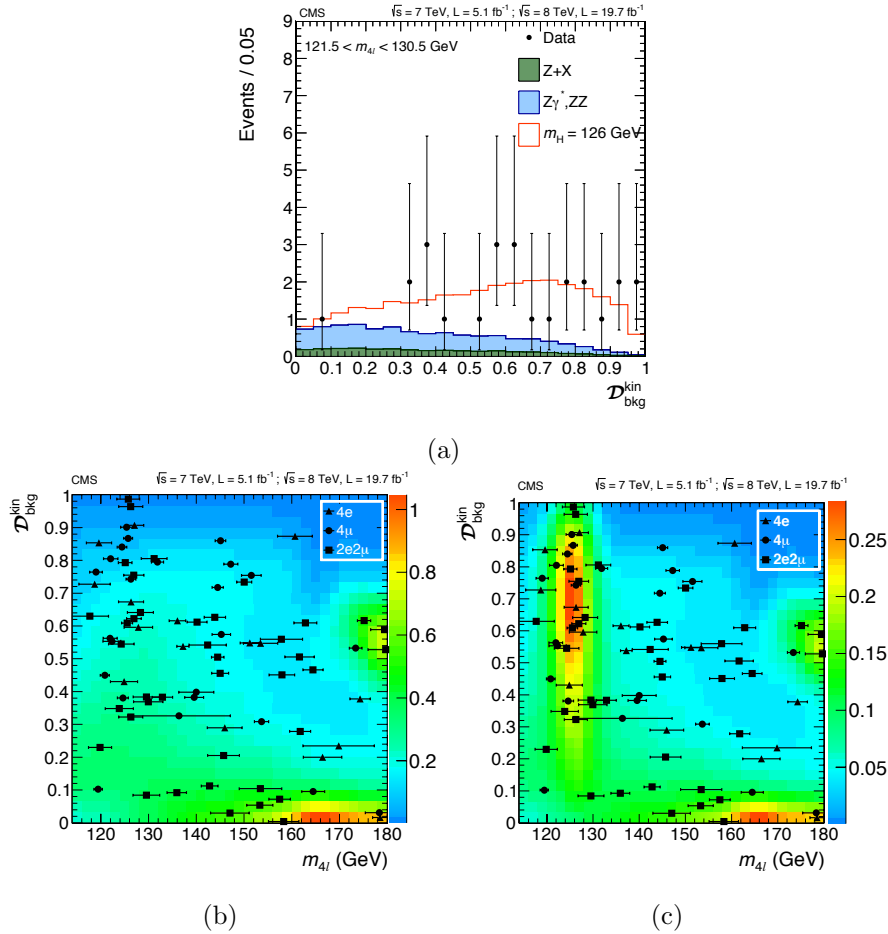


FIGURE 4.5 – Distributions du discriminant cinématique D^{kin} , pour (a) le signal avec $m_H = 126 \text{ GeV}$, les bruits de fond et les données dans la gamme $121.5 < m_{4\ell} < 130.5 \text{ GeV}$, et en fonction de $m_{4\ell}$ pour la gamme $114 < m_{4\ell} < 180 \text{ GeV}$, (b) pour les données et les bruits de fond, et (c) pour les données et la somme des bruits de fond et du signal avec $m_H = 126 \text{ GeV}$.

4.5 Petit historique de la découverte

Avec le recul acquis depuis l'été 2012, la découverte du boson de Higgs peut sembler simple et presque évidente, à tort... Cette section propose un petit rappel historique des résultats obtenus dans la recherche du boson de Higgs lors de 3 points de repère temporels, en juillet 2011 (présentés à l'*International Europhysics Conference on High Energy Physics*, EPS-HEP) [128], décembre 2011 (présentés lors d'un séminaire spécial au CERN) [129] et enfin en juillet 2012 [80, 81, 130], jour de l'annonce de la découverte d'un nouveau boson au CERN. Les distributions et nombres d'événements finaux sont également exposés à la fin de cette section.

Les résultats des limites à 95 % de niveau de confiance sur la section efficace relative à la section efficace du boson de Higgs du modèle standard sont montrés dans la Fig. 4.6 pour la combinaison de toutes les analyses du boson de Higgs en fonction de la masse d'un éventuel boson de Higgs. Lorsque la courbe se situe en dessous de l'unité, la présence d'un boson de Higgs du modèle standard à la masse correspondante est exclue. Notons que l'intégralité de la gamme en masse, à l'exception de celle où se situe le boson de Higgs, a été exclue avant la découverte, ce qui est significatif de la difficulté de la découverte du boson de Higgs à $m_H \approx 125$ GeV.

Les résultats des probabilités de fluctuation de bruit de fond sont montrés dans la Fig. 4.7 pour les mêmes repères temporels. Après un résultat à $\approx 3\sigma$ laissant espérer un boson de Higgs à ≈ 120 GeV ou ≈ 145 GeV dès juillet 2011, les probabilités dépassaient 2σ à ≈ 120 GeV et 3σ à ≈ 125 GeV en décembre 2011, pour laisser place à une probabilité égale à 5σ le jour de l'annonce de la découverte d'un boson de Higgs à ≈ 125 GeV.

En ce qui concerne le canal en quatre leptons, l'observation en juillet 2012 s'est effectuée avec une poignée d'événements [131]. La Fig. 4.8 montre les distributions en $m_{4\ell}$ pour $70 < m_{4\ell} < 180$ GeV au moment où les données ont été révélées dans la région du signal (*unblindées* selon le terme consacré). Les distributions sont montrées pour les données seulement (a), puis en ajoutant les bruits de fond attendus (b), et enfin le signal avec $m_H = 126$ GeV (c).

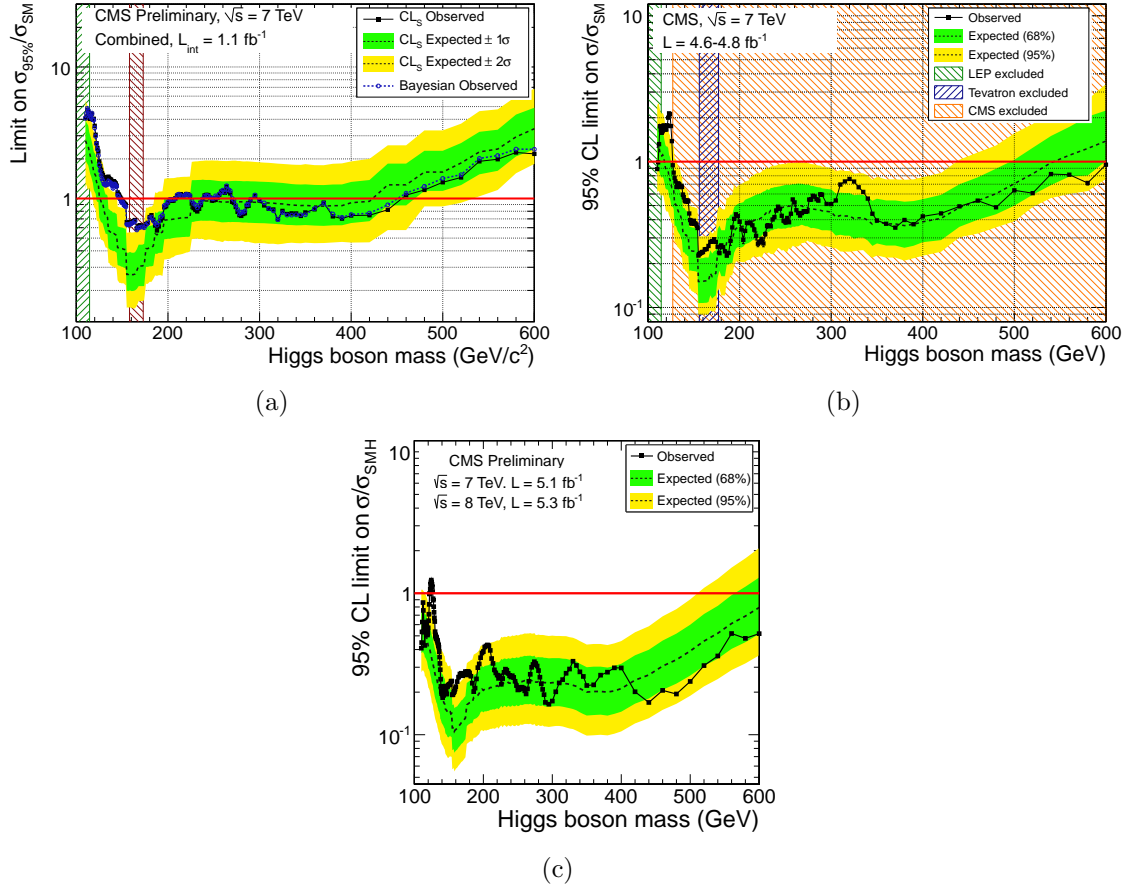


FIGURE 4.6 – Limites à 95 % de niveau de confiance sur la section efficace relative à celle du boson de Higgs du modèle standard, en fonction de la masse d'un éventuel boson de Higgs dans la gamme 110 – 600 GeV, observées (symboles pleins et lignes noires), et attendues (lignes en pointillé). Les bandes vertes et jaunes indiquent les incertitudes à 68 et 95 % sur les limites attendues. Résultats montrés (a) en juillet 2011, (b) en décembre 2011, avec les gammes exclues par le LEP et le Tevatron à ces moments-là, et (c) le 4 juillet 2012.

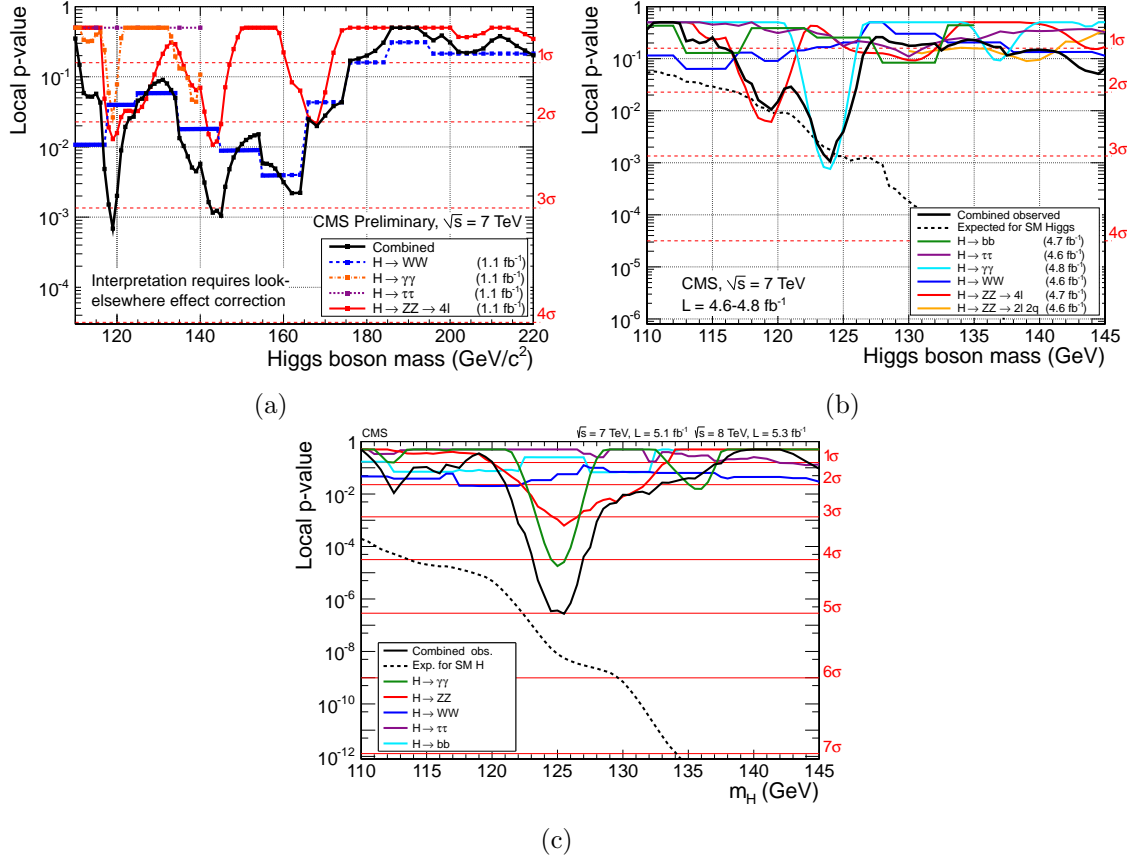
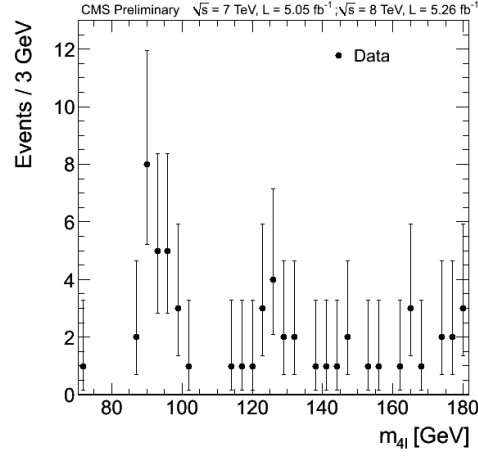
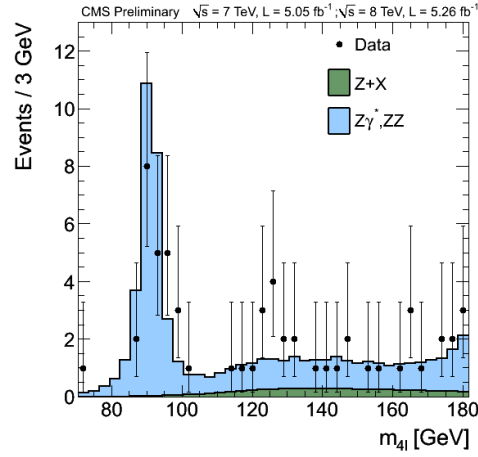


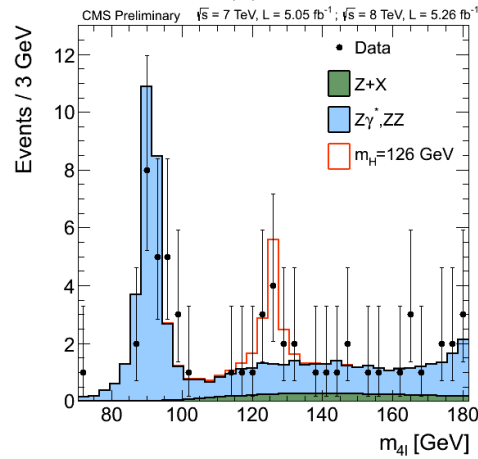
FIGURE 4.7 – Probabilités de fluctuation du bruit de fond observées pour les différents canaux analysés dans la recherche du boson de Higgs (lignes colorées) et leur combinaison (ligne noire), ainsi que celle attendue (ligne en pointillé), (a) en juillet 2011, (b) en décembre 2011, et (c) le 4 juillet 2012.



(a)



(b)



(c)

FIGURE 4.8 – Distributions en $m_{4\ell}$ pour $70 < m_{4\ell} < 180 \text{ GeV}$ au moment où les données ont été *unblindées*, (a) pour les données seulement, (b) en ajoutant les bruits de fond attendus, et (c) en ajoutant le signal avec $m_H = 126 \text{ GeV}$.

Enfin, la Table 4.1 résume les quantités attendues de bruit de fond et de signal à $m_H = 126$ GeV, et le nombre de données observé, obtenus avec la totalité des données enregistrées par CMS et les algorithmes finaux [34]. Les nombres sont donnés pour la gamme en masse $m_{4\ell} > 100$ GeV pour l'ensemble des canaux, et pour la gamme $121.5 < m_{4\ell} < 130.5$ GeV pour chaque canal séparément et leur ensemble. La distribution finale de $m_{4\ell}$ est montrée dans la Fig. 4.9 pour la gamme complète $70 < m_{4\ell} < 1000$ GeV (a) et la gamme à basse masse $70 < m_{4\ell} < 180$ GeV (b). La probabilité de fluctuation de bruit de fond, dont le calcul utilise plusieurs catégories (de prise de données, d'état final, de présence ou non de jets) et une modélisation en trois dimensions, atteint une valeur minimale pour $m_{4\ell} = 125.7$ GeV et correspond à une signification statistique locale égale à 6.8σ , en accord avec la signification attendue à 6.7σ .

Canal	$m_{4\ell} > 100$ GeV		$121.5 < m_{4\ell} < 130.5$ GeV		
	4ℓ	$4e$	$2e2\mu$	4μ	4ℓ
Bruit de fond ZZ	387 ± 31	1.1 ± 0.1	3.2 ± 0.2	2.5 ± 0.2	6.8 ± 0.3
Bruit de fond réductible	22.6 ± 3.6	0.8 ± 0.2	1.3 ± 0.3	0.4 ± 0.2	2.6 ± 0.4
Bruit de fond total	410 ± 31	1.9 ± 0.2	4.6 ± 0.4	2.9 ± 0.2	9.4 ± 0.5
Signal $m_H = 126$ GeV	21.3 ± 1.5	3.4 ± 0.5	9.0 ± 1.1	7.2 ± 0.8	19.6 ± 1.5
Données	470	4	13	8	25

TABLE 4.1 – Nombre d'événements attendus pour le bruit de fond et le signal, et nombre d'événements observés lors de l'analyse du boson de Higgs en quatre leptons, pour $m_{4\ell} > 100$ GeV (ensemble 4ℓ), et $121.5 < m_{4\ell} < 130.5$ GeV (canaux $4e$, $2e2\mu$, 4μ et total 4ℓ).

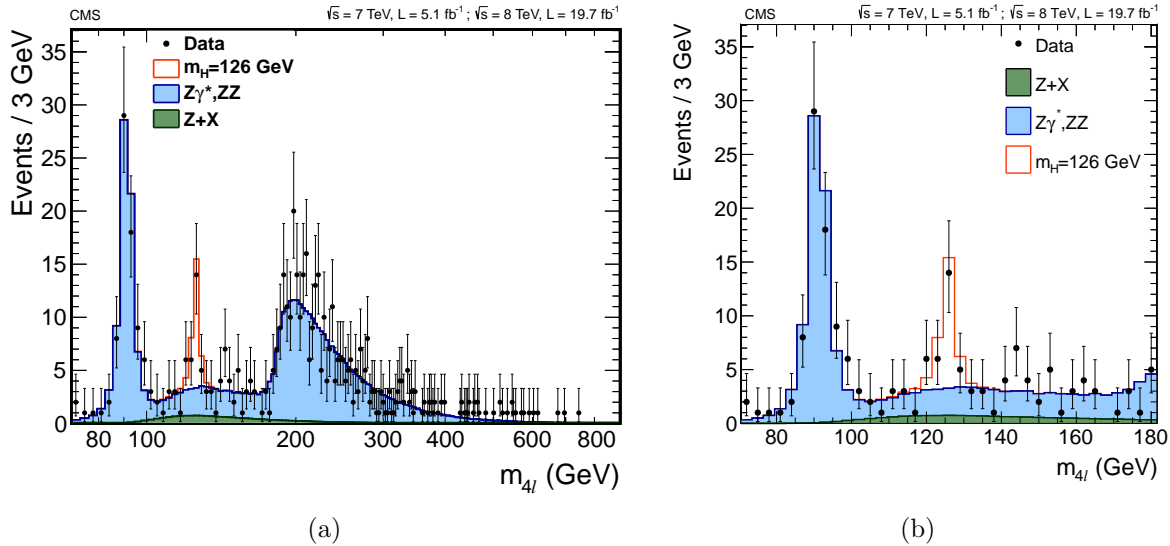


FIGURE 4.9 – Distributions finales en $m_{4\ell}$ pour la gamme (a) complète $70 < m_{4\ell} < 1000$ GeV et (b) à basse masse $70 < m_{4\ell} < 180$ GeV.

4.6 Mesure de la masse

La première mesure des propriétés du boson de Higgs dès sa découverte concerne sa masse. La mesure est basée sur la confrontation des observations expérimentales avec les connaissances théoriques et de la réponse du détecteur, en utilisant la méthode du maximum de vraisemblance avec les événements de données sélectionnés à la fin de l'analyse. Trois niveaux de complexité sont exploités. Le premier, unidimensionnel (1D), est basé sur l'observable $m_{4\ell}$. Le deuxième (2D) y ajoute l'incertitude relative $\sigma_{m_{4\ell}}/m_{4\ell}$ mesurée événement par événement à partir des incertitudes sur les impulsions des leptons. Le dernier (3D) y adjoint le discriminant cinématique D_{kin} , précédemment décrit.

4.6.1 Modélisation de $m_{4\ell}$

La modélisation de $m_{4\ell}$ inclut les densités de probabilité $P(m_{4\ell})$ pour cinq modes de production du signal (gg, VBF, WH, ZH, et $t\bar{t}H$) et trois composantes de bruit de fond ($q\bar{q} \rightarrow ZZ$, $gg \rightarrow ZZ$, et le bruit de fond réductible). Les densités de probabilité pour le bruit de fond sont paramétrées avec des fonctions empiriques issues de la simulation pour ZZ et des régions de contrôle pour le bruit de fond réductible. Pour le signal, la distribution en $m_{4\ell}$ pour une masse m_H donnée est modélisée par une fonction Breit–Wigner (BW) théorique centrée en m_H avec la largeur attendue pour un boson de Higgs du modèle standard (pour $m_H < 400$ GeV) convoluée avec une fonction Crystal Ball (CB) double (explicitée dans la Section 3.4.2.2) décrivant la résolution expérimentale moyenne dans chaque état final. Pour les deux parties de la fonction CB double, seuls les paramètres des queues diffèrent, la résolution σ_{CB} restant la même. Ces distributions sont montrées dans la Fig 4.10 avec leurs paramétrisations pour les trois états finaux $4e$, $2e2\mu$, et 4μ , pour un signal simulé avec $m_H = 126$ GeV. Les résolutions σ_{CB} valent 2.0, 1.6, et 1.2 GeV, et les résolutions effectives σ_{eff} (telles que définies dans la section 3.4.2.2) valent 2.8, 2.3, et 1.7 GeV, respectivement pour les canaux $4e$, $2e2\mu$ et 4μ .

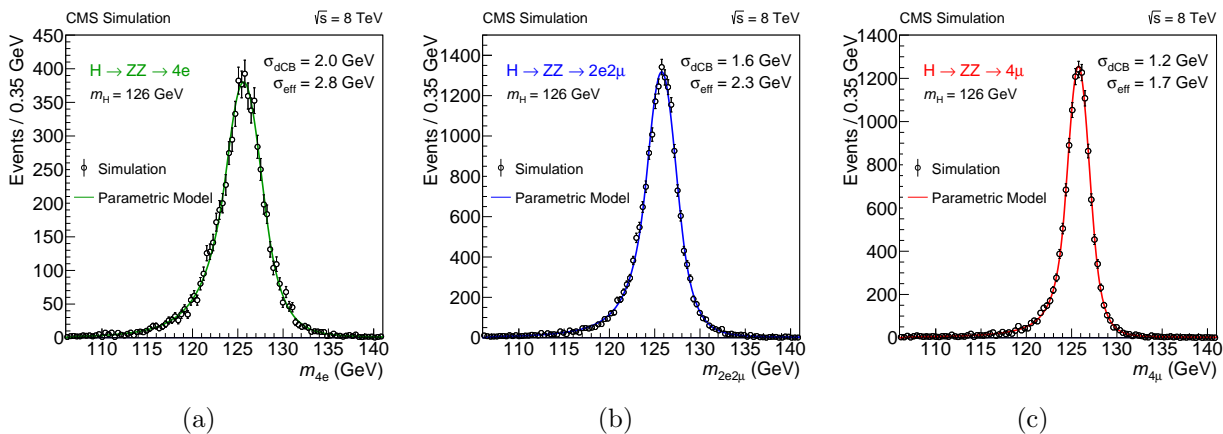


FIGURE 4.10 – Distributions en $m_{4\ell}$ et leurs paramétrisations pour un signal simulé avec $m_H = 126$ GeV pour les canaux (a) $4e$, (b) $2e2\mu$, et (c) 4μ .

4.6.2 Incertitudes sur $m_{4\ell}$ par événement

Le petit nombre d'événements sélectionnés pour l'analyse du boson de Higgs en quatre leptons implique une distribution en $m_{4\ell}$ très variable, en particulier dans la zone de signal : d'hypothétiques autres expériences CMS avec les mêmes conditions de prise de données et d'algorithmie découleraient des distributions en $m_{4\ell}$ très différentes. Cependant il existe un moyen d'approfondir la compréhension de la distribution en $m_{4\ell}$ des événements de données, et ainsi d'affiner la mesure de m_H , par la connaissance des incertitudes sur l'impulsion des leptons. Le concept de base est d'attribuer un poids plus important aux événements les mieux mesurés. L'idée première consistant à classifier les événements en fonction de la qualité de leurs leptons a évolué, progressivement en fonction de la profondeur de la compréhension des leptons, vers la catégorisation ultime, événement par événement. En effet l'incertitude sur l'impulsion est évaluée pour chaque lepton reconstruit dans le détecteur de façon unique, en particulier grâce à la méthode de la régression explicitée dans la Section 3.4 pour les électrons. Cette approche exploitant les incertitudes événement par événement sur $m_{4\ell}$ a un impact significatif sur la mesure de m_H , avec une amélioration attendue pour l'incertitude sur m_H égale à 8 %.

La propagation des incertitudes sur les impulsions des leptons vers l'incertitude sur la masse invariante des quatre leptons requiert une attention particulière. Celle-ci est effectuée par la propagation analytique des incertitudes sur les impulsions des leptons, en négligeant les incertitudes sur les directions, ce qui a été évalué comme équivalent à une somme quadratique des incertitudes des impulsions des leptons². La difficulté de cette propagation réside dans le fait que les distributions des impulsions reconstruites relatives aux impulsions réelles ne sont ni parfaitement symétriques, ni gaussiennes. En particulier pour les électrons, ces deux aspects, quoique significativement réduits par la régression, restent substantiels à bas p_T . Or les incertitudes sur les impulsions des électrons ne prennent en compte ni l'asymétrie des distributions, ni les queues de distribution de façon complète.

L'effet de l'asymétrie est démontré au moyen d'événements de bosons de Higgs simulés pour lesquels les distributions en impulsions générées des leptons (p_{true}) sont distordues à l'aide de distributions à asymétries variables (typiquement la superposition de deux demi-gaussiennes en $(p - p_{\text{true}})/p_{\text{true}}$ piquées à 0 avec des déviations standards différentes). L'asymétrie de la distribution de distorsion est évaluée par la différence entre la valeur moyenne et la valeur maximale de la distribution en $(p - p_{\text{true}})/p_{\text{true}}$. L'erreur sur l'impulsion des leptons, correspondant à la déviation standard effective (telle que définie dans la Section 3.4.2.2) de la distribution en $(p - p_{\text{true}})/p_{\text{true}}$, est propagée pour obtenir l'erreur sur $m_{4\ell}$, $\sigma_{m_{4\ell}}$. La qualité de l'erreur ainsi obtenue, évaluée par la déviation standard effective de la distribution en $(m_{4\ell} - m_H)/\sigma_{m_{4\ell}}$ (escomptée à 1 dans le cas d'erreurs correctement évaluées) est montrée dans la Fig 4.11 en fonction de l'asymétrie de la distribution de distorsion des leptons. L'erreur sur $m_{4\ell}$ est significativement sous-évaluée dans le cas d'une distribu-

2. La propagation des incertitudes prend également en compte les incertitudes sur les photons de FSR éventuels.

tion asymétrique de la distorsion des leptons, avec une ampleur proportionnelle à celle de l'asymétrie, indépendamment de m_H . L'erreur sur m_H obtenue par la propagation des erreurs sur les impulsions des leptons doit par conséquent être corrigée à la hausse.

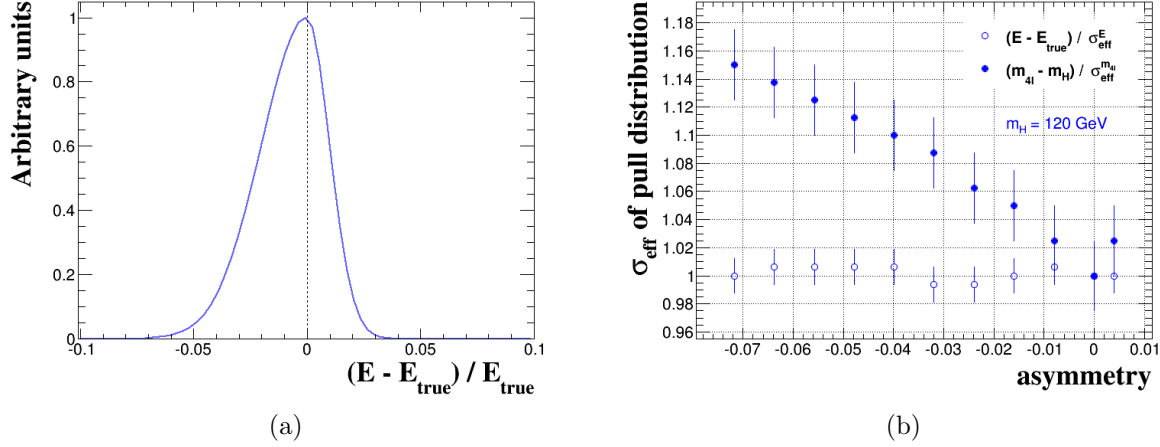


FIGURE 4.11 – (a) Exemple de distribution de distorsion asymétrique en $(p - p_{\text{true}})/p_{\text{true}}$ appliquée à l'impulsion des leptons. (b) Déviation standard effective de la distribution en $(m_{4\ell} - m_H)/\sigma_{m_{4\ell}}$, représentant la qualité de l'erreur $\sigma_{m_{4\ell}}$, en fonction de l'asymétrie de la distribution de distorsion pour $m_H = 120 \text{ GeV}$. La déviation standard de la distribution en $(p - p_{\text{true}})/\sigma_{p_{\text{true}}^{\text{eff}}}$ est également montrée. Les distributions sont similaires pour d'autres valeurs de m_H .

Ces corrections sont mesurées avec des événements $J/\psi \rightarrow \mu^+\mu^-$, $Z \rightarrow \mu^+\mu^-$, et $Z \rightarrow e^+e^-$ dans les données et la simulation, dont les distributions en masse invariante sont ajustées par intervalle de p_T et η à l'aide d'une fonction BW convoluée avec une fonction CB double. Pour un intervalle en p_T et η donné, la résolution de la fonction CB est exprimée par la moyenne des incertitudes sur $m_{2\ell}$ propagées depuis les leptons événement par événement, multipliée par un facteur de correction, paramètre de l'ajustement. Ce facteur de correction multiplicatif est appliqué à l'incertitude de chaque lepton dans les intervalles correspondants en p_T et η . Les valeurs des facteurs de correction ainsi mesurés sont ≈ 1.2 pour les électrons et ≈ 1.1 pour les muons, dans toute la gamme en p_T et η des leptons utilisés dans l'analyse, en accord entre les données et la simulation. Après cette correction, une incertitude de 20 % est attribuée à $\sigma_{m_{4\ell}}$ en comparant les résolutions propagées à partir des leptons et obtenues dans diverses catégories d'événements Z. Par ailleurs les distributions $(m_{4\ell} - m_H)/\sigma_{m_{4\ell}}$ pour les données et les simulations confirment la qualité de $\sigma_{m_{4\ell}}$.

Les densités de probabilité pour $\sigma_{m_{4\ell}}/m_{4\ell}$ sont paramétrées avec des fonctions empiriques issues de la simulation pour le signal et le bruit de fond ZZ et des régions de contrôle pour les bruits de fond réductibles. Les densités de probabilité conditionnelles en deux dimen-

sions pour $m_{4\ell}$ et $\sigma_{m_{4\ell}}/m_{4\ell}$ sont élaborées par la multiplication des densités de probabilité unidimensionnelles, et le remplacement de σ_{CB} par $\sigma_{m_{4\ell}}$ événement par événement dans la modélisation de $m_{4\ell}$ pour le signal. Les paramètres décrivant les queues de la CB sont également multipliés par un facteur égal au rapport $\sigma_{m_{4\ell}}/\sigma_{CB}$.

4.6.3 Discriminant cinématique

Les densités de probabilité en trois dimensions, $m_{4\ell}$, $\sigma_{m_{4\ell}}/m_{4\ell}$ et D_{kin} , sont obtenues en multipliant les densités en deux dimensions précédemment décrites par la densité de probabilité de D_{kin} , négligeant ainsi les corrélations entre D_{kin} et $\sigma_{m_{4\ell}}/m_{4\ell}$, hypothèse validée par la simulation pour le signal et le bruit de fond $q\bar{q} \rightarrow ZZ$.

4.6.4 Incertitudes systématiques

Les incertitudes systématiques principales sur la mesure de m_H proviennent des incertitudes sur l'échelle en impulsion et sur la résolution des leptons, affectant notre connaissance de la forme de la distribution en $m_{4\ell}$. L'incertitude sur l'échelle en impulsion des muons implique une incertitude sur $m_{4\ell}$ égale à 0.1 % dans les canaux 4μ et $2e2\mu$. Concernant les électrons, une dépendance de l'échelle en impulsion valant jusqu'à 0.6 % dans le *barrel* et 1.5 % dans les *endcaps*, est observée et corrigée en fonction de p_T , comme décrit dans la Section 3.4.4. Nonobstant cette correction, cette dépendance mesurée est propagée dans $m_{4\ell}$ et la variation ainsi induite utilisée comme incertitude systématique conservatrice, égale à 0.3 et 0.1 %, pour les canaux $4e$ et $2e2\mu$ respectivement. Les effets des incertitudes sur les résolutions en impulsion des leptons sont considérées en introduisant 20 % d'incertitude sur la résolution de la distribution en $m_{4\ell}$ du signal. Les incertitudes théoriques sur la forme de la distribution du discriminant cinématique sont également considérées, de même que les incertitudes systématiques concernant les normalisations et les formes des différentes contributions de signal et de bruits de fond. L'incertitude systématique totale sur la mesure de m_H est égale à 0.2 %.

4.6.5 Résultats

4.6.5.1 Mesure de m_H avec le canal en quatre leptons dans CMS

La Fig. 4.12 (a) montre les profils de vraisemblance en fonction de l'hypothèse de m_H pour les trois niveaux de complexité des modèles exploités pour la mesure. Les valeurs des mesures de m_H , correspondant aux minimums des profils, sont indiquées dans la Table 4.2. Celles-ci sont très proches, leurs incertitudes diminuant en fonction de la complexité du modèle et donc de l'information exploitée. Le gain sur l'incertitude de la mesure de m_H apporté par l'exploitation de $\sigma_{m_{4\ell}}/m_{4\ell}$ événement par événement est égal à ≈ 6 %, le gain supplémentaire apporté par l'exploitation du discriminant cinématique est égal à ≈ 5 %.

La Fig. 4.12 (b) montre les résultats pour le modèle le plus complet (3D) pour chacun des canaux $4e$, $2e2\mu$, et 4μ et et leur combinaison. Les valeurs des mesures correspondantes sont statistiquement compatibles, et sont indiquées dans la Table 4.3.

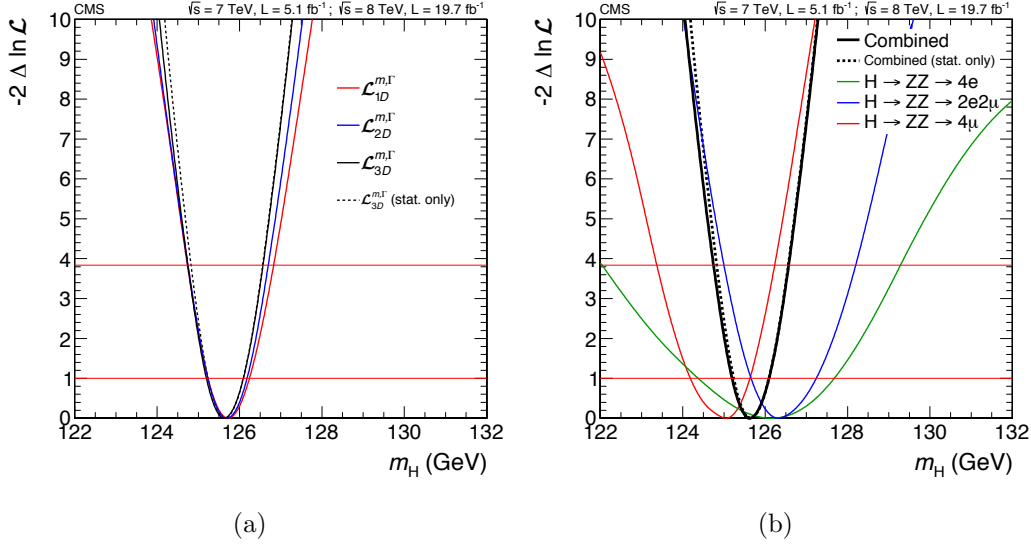


FIGURE 4.12 – Profils de vraisemblance en fonction de l’hypothèse de m_H : (a) pour les trois niveaux de complexité de la mesure, 1D avec $m_{4\ell}$ (courbe rouge), 2D avec de plus les incertitudes $\sigma_{m_{4\ell}}/m_{4\ell}$ événement par événement (courbe bleue), et 3D en ajoutant le discriminant cinématique D_{kin} (courbe noire, la courbe en pointillé incluant uniquement les erreurs statistiques); (b) pour la méthode 3D, pour chacun des canaux, $4e$ (courbe verte), $2e2\mu$ (courbe bleue), et 4μ (courbe rouge), et leur combinaison (courbe noire, la courbe en pointillé incluant uniquement les erreurs statistiques).

Canal	m_H 1D (GeV)	m_H 2D (GeV)	m_H 3D (GeV)
4ℓ	$125.72^{+0.52}_{-0.49}$	$125.69^{+0.50}_{-0.45}$	$125.63^{+0.47}_{-0.39}$ (stat.) $^{+0.09}_{-0.18}$ (syst.)

TABLE 4.2 – Mesure de m_H pour chacune des trois méthodes : 1D avec $m_{4\ell}$, 2D avec de plus les incertitudes $\sigma_{m_{4\ell}}/m_{4\ell}$ événement par événement, et 3D en ajoutant le discriminant cinématique D_{kin} . Les incertitudes sont indiquées de façon globale pour les méthode 1D et 2D, et séparément pour la méthode 3D.

Canal	m_H (GeV)
4e	$126.2^{+1.5}_{-1.8}$
2e2 μ	$126.3^{+0.9}_{-0.7}$
4 μ	$125.1^{+0.6}_{-0.9}$
4 ℓ	$125.6 \pm 0.4(\text{stat.}) \pm 0.2(\text{syst.})$

TABLE 4.3 – Mesure de m_H pour la méthode 3D pour chacun des canaux, 4e, 2e2 μ , et 4 μ , et leur combinaison. Les incertitudes sont indiquées de façon globale pour les canaux individuels, et séparément pour la combinaison.

La mesure finale de la masse du boson de Higgs dans le canal en quatre leptons pour CMS est la suivante³ :

$$m_H = 125.6 \pm 0.4 \text{ (stat.)} \pm 0.2 \text{ (syst.) GeV.}$$

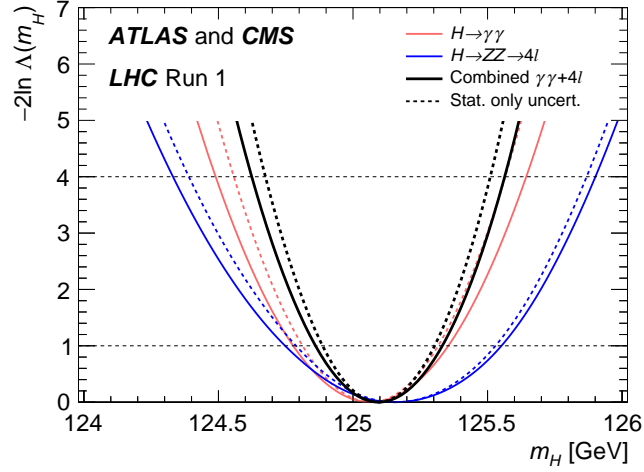
4.6.5.2 Combinaison ultime

Ce résultat a été exploité dans la combinaison finale des canaux les plus sensibles, en quatre leptons et en deux photons, des expériences ATLAS et CMS [132]. Les profils de vraisemblance en fonction de l'hypothèse de m_H sont montrés dans la Fig. 4.13 (a) pour les expériences ATLAS et CMS combinées, pour le canal en deux photons, le canal en quatre leptons, et la combinaison finale des canaux. Les mesures individuelles et la combinaison finale sont montrées dans la Fig. 4.13 (b).

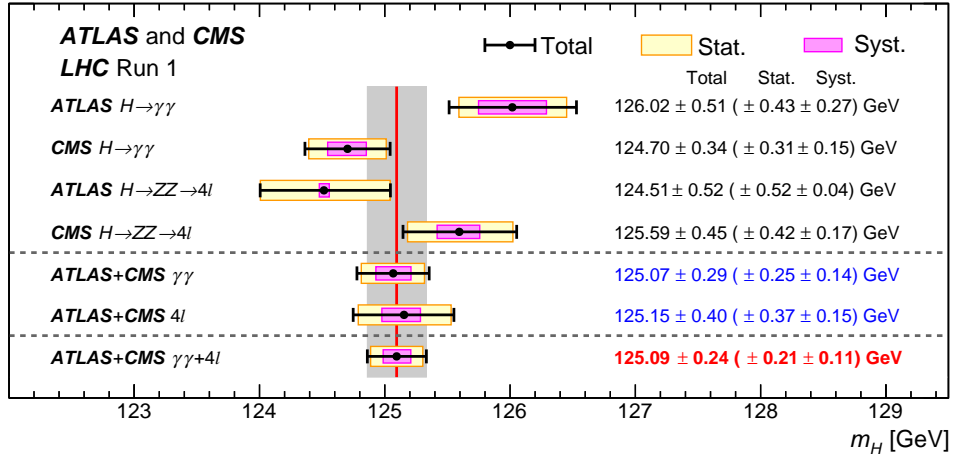
La mesure finale de la masse du boson de Higgs pour la combinaison ultime est la suivante :

$$m_H = 125.09 \pm 0.24 \text{ (}\pm 0.21 \text{ (stat.)} \pm 0.11 \text{ (syst.)) GeV.}$$

3. Entre les valeurs données dans la Table 4.2 et la mesure finale reportée ici et dans la Table 4.3, le profil de vraisemblance a été légèrement affiné, et par conséquent les valeurs des incertitudes très légèrement modifiées.



(a)



(b)

FIGURE 4.13 – (a) Profils de vraisemblance en fonction de l’hypothèse de m_H pour les expériences ATLAS et CMS combinées, pour le canal en deux photons (courbes rouges), le canal en quatre leptons (courbes bleues), et la combinaison finale des canaux (courbes noires), les courbes en pointillé incluant uniquement les erreurs statistiques. (b) Valeurs des mesures de m_H , individuelles et pour la combinaison finale.

En un peu plus de 3 années d'exploitation intensive des données du LHC, ce boson de Higgs si élu­si­f est devenu non seulement une réalité tangible, mais a subi un interrogatoire si poussé que ses caractéristiques sont dorénavant connues de façon presque intime, en particulier sa masse, 125.09 GeV, avec une incertitude de 0.2 %. Mesurer, cependant, est l'aboutissement d'une longue chaîne lors de laquelle il a été nécessaire de reconstruire les précieux événements, et auparavant de déclencher leur enregistrement, après les avoir produits. La réalisation et le succès d'une telle aventure ont été possibles uniquement par une préparation longue et minutieuse sur tous les fronts, ainsi qu'une attention de tous les instants lors de la prise de données et l'élaboration des résultats.

Conclusion

Tout au long de ce manuscrit, les ingrédients nécessaires à la découverte du boson de Higgs et la mesure de sa masse dans l'expérience CMS auprès du LHC ont été développés, pour le canal majeur de désintégration en quatre leptons, avec une emphase sur la reconstruction et l'identification des électrons. Les performances de chaque étape, de la production à l'analyse finale, ont été exposées. J'espère avoir convaincu le lecteur de l'importance du rôle des électrons dans ce canal, ainsi que de la nécessité de finesse et de précision sur la connaissance des leptons dans une telle analyse, qui peut paraître à première vue très simple. Selon les derniers résultats, la masse du boson de Higgs est estimée avec une précision de 0.2% en combinant les analyses et les données de CMS et d'ATLAS. Les autres propriétés du boson de Higgs sont également mesurées en combinant les deux expériences, et les résultats récents, non abordés dans ce manuscrit, se trouvent dans la Ref. [133], montrant la connaissance maintenant très approfondie de ce fameux boson.

La recherche et la découverte du boson de Higgs se sont révélées une aventure extraordinaire, scientifiquement et humainement. Un long parcours, semé d'embûches, d'espoirs, de doutes, et de joies. Une illustration notoire de la collaboration efficace dont ont fait montre les acteurs, à laquelle j'ai eu la chance de contribuer. Les trois événements publics mentionnés dans le dernier chapitre ont été tels des jalons dans la conquête. Depuis EPS à Grenoble en juillet 2011, j'écrivais : "Il y a des *bumps* par ci par là, tout le monde met des limites, et le Higgs est à 140 GeV (interprétation personnelle)"... En décembre 2011, la fenêtre de possibilité de la masse du boson de Higgs se réduisant drastiquement (et avec elle les probabilités de se tromper), d'aucuns pariaient sur une masse à ≈ 119 GeV (disons qu'il est plutôt bien venu que l'intuition ne soit pas l'unique composante d'une découverte scientifique). Peu avant l'annonce officielle du 4 juillet 2012, l'*unblinding* a constitué une étape cruciale : le premier regard sur l'émergence du pic au dessus du bruit de fond dans la figure représentant la masse invariante des quatre leptons laisse un souvenir vif et mémorable. Quant au 4 juillet 2012, l'expérience émotionnelle et humaine était à son comble. Voici ce que j'écrivais aux alentours de 5h20 du matin ce jour-là : "C'est la folie douce devant l'amphi, les gens dorment par terre, il y a déjà au moins 300 personnes qui attendent, et certains dont je tairais le nom étaient déjà là à minuit !! Great moment !... Il y a une alarme incendie depuis 10 min mais PERSONNE ne bouge ! Les pompiers ont essayé de nous évacuer, pas moyen !". Du transparent illustrant la découverte, judicieusement montré sans aucun commentaire par Joe Incandela (porte-parole de CMS), a émané un silence, d'une essence palpable, laissant dans son sillage

un frisson généralisé.

L'histoire du boson de Higgs est cependant loin d'être achevée, avec les données à 13 TeV dans le centre de masse collectées au moment même où j'écris ces mots et le futur à très haute luminosité du LHC. Mais c'est ici que je m'arrête, en remerciant le lecteur pour son intérêt. Qu'ai-je fait durant ces années ? Je pense avoir apporté quelques pierres à cet édifice, ce qui en soit est une immense satisfaction.

Bibliographie

Chapitre 1 : Produire

- [1] F. Englert and R. Brout. Broken Symmetry and the Mass of Gauge Vector Mesons. *Phys. Rev. Lett.*, 13 :321–323, 1964.
- [2] Peter W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett.*, 12 :132–133, 1964.
- [3] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13 :508–509, 1964.
- [4] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble. Global Conservation Laws and Massless Particles. *Phys. Rev. Lett.*, 13 :585–587, 1964.
- [5] Peter W. Higgs. Spontaneous Symmetry Breakdown without Massless Bosons. *Phys. Rev.*, 145 :1156–1163, 1966.
- [6] T. W. B. Kibble. Symmetry breaking in nonAbelian gauge theories. *Phys. Rev.*, 155 :1554–1561, 1967.
- [7] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22 :579–588, 1961.
- [8] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19 :1264–1266, 1967.
- [9] Abdus Salam. Weak and Electromagnetic Interactions. *Conf. Proc.*, C680519 :367–377, 1968.
- [10] David J. Gross and Frank Wilczek. Ultraviolet Behavior of Nonabelian Gauge Theories. *Phys. Rev. Lett.*, 30 :1343–1346, 1973.
- [11] H. David Politzer. Reliable Perturbative Results for Strong Interactions? *Phys. Rev. Lett.*, 30 :1346–1349, 1973.
- [12] Yoichiro Nambu and G. Jona-Lasinio. Dynamical Model of Elementary Particles Based on an Analogy with Superconductivity. 1. *Phys. Rev.*, 122 :345–358, 1961.
- [13] Murray Gell-Mann and M Levy. The axial vector current in beta decay. *Nuovo Cim.*, 16 :705, 1960.
- [14] John M. Cornwall, David N. Levin, and George Tiktopoulos. Uniqueness of spontaneously broken gauge theories. *Phys. Rev. Lett.*, 30 :1268–1270, 1973. [Erratum : *Phys. Rev. Lett.* 31,572(1973)].

- [15] John M. Cornwall, David N. Levin, and George Tiktopoulos. Derivation of Gauge Invariance from High-Energy Unitarity Bounds on the s Matrix. *Phys. Rev.*, D10 :1145, 1974. [Erratum : *Phys. Rev.*D11,972(1975)].
- [16] C. H. Llewellyn Smith. High-Energy Behavior and Gauge Symmetry. *Phys. Lett.*, B46 :233–236, 1973.
- [17] Benjamin W. Lee, C. Quigg, and H. B. Thacker. Weak Interactions at Very High-Energies : The Role of the Higgs Boson Mass. *Phys. Rev.*, D16 :1519, 1977.
- [18] ALEPH, CDF, D0, DELPHI, L3, OPAL, and SLD Collaborations, the LEP Electroweak Working Group, the Tevatron Electroweak Working Group, and the SLD Electroweak and Heavy Flavour Groups. Precision electroweak measurements and constraints on the standard model. *PH-EP-2010-095*, 2010.
At this time, the most up-to-date Higgs boson mass constraints come from :
https://lepewwg.web.cern.ch/LEPEWWG/plots/lep2_2013.
- [19] R. Barate et al. Search for the standard model Higgs boson at LEP. *Phys. Lett.*, B565 :61–75, 2003.
- [20] T. Aaltonen et al. Combination of Tevatron searches for the standard model Higgs boson in the W+W- decay mode. *Phys. Rev. Lett.*, 104 :061802, 2010.
- [21] J R Andersen et al. Handbook of LHC Higgs Cross Sections : 3. Higgs Properties. *CERN-2013-004, FERMILAB-CONF-13-667-T*, 2013.
- [22] S. Chatrchyan et al. The CMS experiment at the CERN LHC. *JINST*, 3 :S08004, 2008.
- [23] Serguei Chatrchyan et al. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *JINST*, 9(10) :P10009, 2014.
- [24] N. Almeida et al. Data filtering in the readout of the CMS electromagnetic calorimeter. *JINST*, 3 :P02011, 2008.
- [25] P. Adzic et al. Energy resolution of the barrel of the CMS electromagnetic calorimeter. *JINST*, 2 :P04004, 2007.

Chapitre 2 : Déclencher

- [26] S. Dasu et al. CMS. The TriDAS project. Technical design report, vol. 1 : The trigger systems. *CERN-LHCC-2000-038*, 2000.
- [27] Vardan Khachatryan et al. The CMS Trigger. *To be submitted to JINST, CMS PAPER TRG-12-001 (CMS internal for now)*.
- [28] P. Sphicas. CMS : The TriDAS project. Technical design report, Vol. 2 : Data acquisition and high-level trigger. 2002.
- [29] Stéphanie Beauceron. The CMS High Level Trigger. *PoS, ICHEP2012* :505, 2013.
- [30] Alexandre Zabi. The electron and photon trigger of the CMS experiment. *CMS-CR-2013-256, CHEF2013*.

- [31] Pascal Paganini. Genèse, mise au point et performances du système de déclenchement du calorimètre électromagnétique de CMS. *Thèse d'habilitation*, 2011.
<https://11rweb.in2p3.fr/~paganini/hdr.pdf>.
- [32] Nadir Daci. Sélection des électrons et recherche du boson de Higgs se désintégrant en paires de leptons tau avec l'expérience CMS au LHC. *Thèse de doctorat, CMS-TS-2014-037, CERN-THESIS-2013-344*, 2013.
- [33] The Higgs PAG. Trigger strategies for Higgs searches. *CMS Internal, CMS AN-2011/065*, 2011.
- [34] Serguei Chatrchyan et al. Measurement of the properties of a Higgs boson in the four-lepton final state. *Phys.Rev.*, D89(9) :092007, 2014.
- [35] Twiki page : Egamma Working Points. *CMS internal*, 2013.
<https://twiki.cern.ch/twiki/bin/viewauth/CMS/EgammaWorkingPointsv3>.
- [36] Web page : CMS RunSummary Run 208686. *CMS internal*.
<https://cmswbm.web.cern.ch/cmswbm/cmsdb/servlet/RunSummary?RUN=208686&SUBMIT=Submit>.
- [37] Vardan Khachatryan et al. Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 01 :080, 2011.
- [38] Pauline Soulet. Mesure de l'efficacité de déclenchement des événements $Higgs \rightarrow ZZ^* \rightarrow 4e$. *Rapport de stage de M2, École centrale de Lyon*, 2012.
<http://11r.in2p3.fr/~baffioni/students/Pauline.pdf>.

Chapitre 3 : Reconstruire

- [39] G.L. Bayatian et al. CMS physics : Technical design report. *CERN-LHCC-2006-001*, 2006.
- [40] S. Baffioni, C. Charlot, F. Ferri, D. Futyan, P. Meridiani, I. Puljak, C. Rovelli, R. Salerno, and Y. Sirois. Electron reconstruction in CMS. *Eur. Phys. J.*, C49 :1099–1116, 2007.
- [41] Claude Charlot. La reconstruction des électrons dans CMS. *Thèse d'habilitation*, 2007.
https://11rweb.in2p3.fr/~charlot/Hdr/electrons_hdr.pdf.
- [42] Vardan Khachatryan et al. Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV. *JINST*, 10(06) :P06005, 2015.
- [43] S. Chatrchyan et al. Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. *CMS Physics Analysis Summary, CMS-PAS-PFT-09-001*, 2009.
- [44] S. Chatrchyan et al. Commissioning of the Particle-Flow Event Reconstruction with the first LHC collisions recorded in the CMS detector. *CMS Physics Analysis Summary, CMS-PAS-PFT-10-001*, 2010.

- [45] S. Chatrchyan et al. CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update. *CMS Physics Analysis Summary, CMS-PAS-LUM-13-001*, 2013.
- [46] Serguei Chatrchyan et al. Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 7$ TeV. *JINST*, 8 :P09009, 2013.
- [47] Serguei Chatrchyan et al. Alignment of the CMS tracker with LHC and cosmic ray data. *JINST*, 9 :P06009, 2014.
- [48] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07 :079, 2014.
- [49] Torbjorn Sjöstrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05 :026, 2006.
- [50] Serguei Chatrchyan et al. Study of the underlying event at forward rapidity in pp collisions at $\sqrt{s} = 0.9, 2.76$, and 7 TeV. *JHEP*, 04 :072, 2013.
- [51] S. Agostinelli et al. GEANT4 : A Simulation toolkit. *Nucl. Instrum. Meth.*, A506 :250–303, 2003.
- [52] John Allison et al. Geant4 developments and applications. *IEEE Trans. Nucl. Sci.*, 53 :270, 2006.
- [53] Wolfgang Adam, R. Frühwirth, Are Strandlie, and T. Todor. Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC. *CMS-NOTE-2005-001*, 2005.
- [54] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov. Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at LHC. *J. Phys. G*, 31 :N9, 2005.
- [55] Andreas Hocker et al. TMVA - Toolkit for Multivariate Data Analysis. *PoS, ACAT* :040, 2007.
- [56] Vardan Khachatryan et al. Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV. *JINST*, 10(08) :P08010, 2015.
- [57] Arabella Martelli. First validation of the new geometry in the simulation. *CMS Internal talk*, 2014. https://indico.cern.ch/event/333019/contribution/4/attachments/649433/893103/AMartelli.PPD_30Jul14.pdf.
- [58] Clémentine Broutin. Electron Measurements and Search for Higgs Bosons in Multi-Lepton Channels with the CMS Experiment at LHC. *Thèse de doctorat, CMS-TS-2011-048, CERN-THESIS-2011-108*, 2011.
- [59] Serguei Chatrchyan et al. Measurement of the electron charge asymmetry in inclusive W production in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Rev. Lett.*, 109 :111806, 2012.
- [60] Serguei Chatrchyan et al. Search for new physics in events with same-sign dileptons and jets in pp collisions at $\sqrt{s} = 8$ TeV. *JHEP*, 01 :163, 2014.

- [61] S. Baffioni et al. Electron Charge Identification using 8 TeV data. *CMS Internal, CMS AN-2014/164*, 2014.
- [62] M. Anfreville et al. Laser monitoring system for the CMS lead tungstate crystal calorimeter. *Nucl. Instrum. Meth.*, A594 :292–320, 2008.
- [63] Li-Yuan Zhang, D. Bailleux, A. Bornheim, Ke-Jun Zhu, and Ren-Yuan Zhu. Performance of the monitoring light source for the CMS lead tungstate crystal calorimeter. *IEEE Trans. Nucl. Sci.*, 52 :1123–1130, 2005.
- [64] Jerome H. Friedman. Greedy Function Approximation : A Gradient Boosting Machine. *Ann. of Stat.*, 29 :1189, 2001.
- [65] F. Beaudette et al. Electron momentum determination using boosted regression trees. *CMS Internal, CMS AN-2013/209*, 2013.
- [66] S. Catani, Yuri L. Dokshitzer, M.H. Seymour, and B.R. Webber. Longitudinally invariant k_T clustering algorithms for hadron hadron collisions. *Nucl. Phys.*, B406 :187–224, 1993.
- [67] Stephen D. Ellis and Davison E. Soper. Successive combination jet algorithm for hadron collisions. *Phys. Rev.*, D48 :3160–3166, 1993.
- [68] M. Oreglia. A study of the reactions $\psi' \rightarrow \gamma\gamma\psi$. *PhD. Thesis, SLAC-0236*, 1980.
- [69] K. A. Olive et al. Review of Particle Physics. *Chin. Phys.*, C38 :090001, 2014.
- [70] GEANT4 Collaboration. GEANT4 10.0 Release Notes. 2013.
<http://geant4.cern.ch/support/ReleaseNotes4.10.0>.
- [71] Muriel Pivk and Francois R. Le Diberder. SPlot : A Statistical tool to unfold data distributions. *Nucl. Instrum. Meth.*, A555 :356–369, 2005.
- [72] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Phys. Lett.*, B659 :119–126, 2008.
- [73] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Catchment Area of Jets. *JHEP*, 04 :005, 2008.
- [74] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet User Manual. *Eur. Phys. J.*, C72 :1896, 2012.

Chapitre 4 : Mesurer

- [75] N. Amapane et al. Measurement of the production and decay of a Higgs boson in the four-lepton final state. *CMS Internal, CMS AN-2013/108*, 2013.
- [76] Mykhailo Dalchenko. Measurements of the Higgs boson mass and width in the four-lepton final state and electron reconstruction in the CMS experiment at the LHC. *PhD. Thesis, CERN-THESIS-2014-250*, 2014.
- [77] Vardan Khachatryan et al. Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV. *Eur. Phys. J.*, C75(5) :212, 2015.

- [78] G. L. Bayatian et al. CMS technical design report, volume II : Physics performance. *J. Phys.*, G34 :995–1579, 2007.
- [79] S. Baffioni, C. Charlot, F. Ferri, N. Godinovic, P. Meridiani, I. Puljak, R. Salerno, and Y. Sirois. Discovery potential for the SM Higgs boson in the $H \rightarrow ZZ^{(*)} \rightarrow e^+e^-e^+e^-$ decay channel. *J. Phys.*, G34 :N23–N46, 2007.
- [80] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716 :30–61, 2012.
- [81] Serguei Chatrchyan et al. Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV. *JHEP*, 06 :081, 2013.
- [82] Serguei Chatrchyan et al. Electron performance with 19.6 fb⁻¹ of data collected at $\sqrt{s} = 8$ TeV with the CMS detector. *CMS Detector Performance Summary, CERN-CMS-DP-2013-003*, 2013.
- [83] Vardan Khachatryan et al. Electron commissioning results at $\sqrt{s} = 7$ TeV. *CMS Detector Performance Summary, CERN-CMS-DP-2011-003*, 2011.
- [84] S. Chatrchyan et al. Electron reconstruction and identification at $\sqrt{s} = 7$ TeV. *CMS Physics Analysis Summary, CMS-PAS-EGM-10-004*, 2010.
- [85] S. Baffioni et al. Electron Momentum Determination Commissioning from 7 TeV Data. *CMS Internal, CMS AN-2010/472*, 2010.
- [86] S. Baffioni et al. Electron Efficiency Measurement from 7 TeV Data. *CMS Internal, CMS AN-2010/471*, 2010.
- [87] S. Baffioni et al. Electron Track Seeding Commissioning from 7 TeV Data. *CMS Internal, CMS AN-2010/470*, 2010.
- [88] S. Baffioni et al. Electron Charge Determination Commissioning from 7 TeV Data. *CMS Internal, CMS AN-2010/468*, 2010.
- [89] S. Baffioni et al. Electron Commissioning in CMS from first 7 TeV Minimum Bias Data. *CMS Internal, CMS AN-2010/235*, 2010.
- [90] S. Baffioni et al. Electron Commissioning from First CMS Data. *CMS Internal, CMS AN-2010/037*, 2010.
- [91] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations : the POWHEG method. *JHEP*, 11 :070, 2007.
- [92] E. Bagnaschi, G. Degrandi, P. Slavich, and A. Vicini. Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM. *JHEP*, 02 :088, 2012.
- [93] Paolo Nason and Carlo Oleari. NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG. *JHEP*, 02 :037, 2010.
- [94] D. de Florian, G. Ferrera, M. Grazzini, and D. Tommasini. Higgs boson production at the LHC : transverse momentum resummation effects in the $H \rightarrow 2\gamma$, $H \rightarrow WW \rightarrow \nu\nu\nu$ and $H \rightarrow ZZ \rightarrow 4l$ decay modes. *JHEP*, 06 :132, 2012.
- [95] Charalampos Anastasiou, Radja Boughezal, and Frank Petriello. Mixed QCD-electroweak corrections to Higgs boson production in gluon fusion. *JHEP*, 04 :003, 2009.

- [96] Daniel de Florian and Massimiliano Grazzini. Higgs production through gluon fusion : Updated cross sections at the Tevatron and the LHC. *Phys. Lett.*, B674 :291–294, 2009.
- [97] Julien Baglio and Abdelhak Djouadi. Higgs production at the LHC. *JHEP*, 03 :055, 2011.
- [98] S. Dittmaier et al. Handbook of LHC Higgs Cross Sections : 1. Inclusive Observables. *CERN-2011-002*, 2011.
- [99] A. Djouadi, M. Spira, and P. M. Zerwas. Production of Higgs bosons in proton colliders : QCD corrections. *Phys. Lett.*, B264 :440–446, 1991.
- [100] S. Dawson. Radiative corrections to Higgs boson production. *Nucl. Phys.*, B359 :283–300, 1991.
- [101] M. Spira, A. Djouadi, D. Graudenz, and P. M. Zerwas. Higgs boson production at the LHC. *Nucl. Phys.*, B453 :17–82, 1995.
- [102] Robert V. Harlander and William B. Kilgore. Next-to-next-to-leading order Higgs production at hadron colliders. *Phys. Rev. Lett.*, 88 :201801, 2002.
- [103] Charalampos Anastasiou and Kirill Melnikov. Higgs boson production at hadron colliders in NNLO QCD. *Nucl. Phys.*, B646 :220–256, 2002.
- [104] V. Ravindran, J. Smith, and W. L. van Neerven. NNLO corrections to the total cross-section for Higgs boson production in hadron hadron collisions. *Nucl. Phys.*, B665 :325–366, 2003.
- [105] Stefano Catani, Daniel de Florian, Massimiliano Grazzini, and Paolo Nason. Soft gluon resummation for Higgs boson production at hadron colliders. *JHEP*, 07 :028, 2003.
- [106] Stefano Actis, Giampiero Passarino, Christian Sturm, and Sandro Uccirati. NLO Electroweak Corrections to Higgs Boson Production at Hadron Colliders. *Phys. Lett.*, B670 :12–17, 2008.
- [107] M. Ciccolini, Ansgar Denner, and S. Dittmaier. Strong and electroweak corrections to the production of Higgs + 2jets via weak interactions at the LHC. *Phys. Rev. Lett.*, 99 :161803, 2007.
- [108] Mariano Ciccolini, Ansgar Denner, and Stefan Dittmaier. Electroweak and QCD corrections to Higgs production via vector-boson fusion at the LHC. *Phys. Rev.*, D77 :013002, 2008.
- [109] T. Figy, C. Oleari, and D. Zeppenfeld. Next-to-leading order jet distributions for Higgs boson production via weak boson fusion. *Phys. Rev.*, D68 :073005, 2003.
- [110] K. Arnold et al. VBFNLO : A Parton level Monte Carlo for processes with electroweak bosons. *Comput. Phys. Commun.*, 180 :1661–1670, 2009.
- [111] Paolo Bolzoni, Fabio Maltoni, Sven-Olaf Moch, and Marco Zaro. Higgs production via vector-boson fusion at NNLO in QCD. *Phys. Rev. Lett.*, 105 :011801, 2010.
- [112] Giampiero Passarino. Higgs Interference Effects in $gg \rightarrow ZZ$ and their Uncertainty. *JHEP*, 08 :146, 2012.

- [113] Yanyan Gao, Andrei V. Gritsan, Zijin Guo, Kirill Melnikov, Markus Schulze, and Nhan V. Tran. Spin determination of single-produced resonances at hadron colliders. *Phys. Rev.*, D81 :075022, 2010.
- [114] Sara Bolognesi, Yanyan Gao, Andrei V. Gritsan, Kirill Melnikov, Markus Schulze, Nhan V. Tran, and Andrew Whitbeck. On the spin and parity of a single-produced resonance at the LHC. *Phys. Rev.*, D86 :095031, 2012.
- [115] Ian Anderson et al. Constraining anomalous HVV interactions at proton and lepton colliders. *Phys. Rev.*, D89(3) :035007, 2014.
- [116] Tom Melia, Paolo Nason, Raoul Rontsch, and Giulia Zanderighi. W+W-, WZ and ZZ production in the POWHEG BOX. *JHEP*, 11 :078, 2011.
- [117] T. Binoth, N. Kauer, and P. Mertsch. Gluon-induced QCD corrections to $pp \rightarrow ZZ \rightarrow l\bar{l}l'\bar{l}'$. In *Proceedings, 16th International Workshop on Deep Inelastic Scattering and Related Subjects (DIS 2008)*, page 142, 2008.
- [118] John M. Campbell and R. K. Ellis. MCFM for the Tevatron and the LHC. *Nucl. Phys. Proc. Suppl.*, 205-206 :10–15, 2010.
- [119] John M. Campbell and R. Keith Ellis. An Update on vector boson pair production at hadron colliders. *Phys. Rev.*, D60 :113006, 1999.
- [120] John M. Campbell, R. Keith Ellis, and Ciaran Williams. Vector boson pair production at the LHC. *JHEP*, 07 :018, 2011.
- [121] Serguei Chatrchyan et al. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV. *JINST*, 7 :P10002, 2012.
- [122] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, Jon Pumplin, and C. P. Yuan. New parton distributions for collider physics. *Phys. Rev.*, D82 :074024, 2010.
- [123] Sergey Alekhin et al. The PDF4LHC Working Group Interim Report. 2011.
- [124] Michiel Botje et al. The PDF4LHC Working Group Interim Recommendations. 2011.
- [125] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J.*, C63 :189–285, 2009.
- [126] Richard D. Ball, Valerio Bertone, Francesco Cerutti, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Jose I. Latorre, Juan Rojo, and Maria Ubiali. Impact of Heavy Quark Masses on Parton Distributions and LHC Phenomenology. *Nucl. Phys.*, B849 :296–363, 2011.
- [127] Vardan Khachatryan et al. Absolute Calibration of Luminosity Measurement at CMS : Summer 2011 Update. *CMS Physics Analysis Summary, CMS-PAS-EWK-11-001*, 2011.
- [128] Vardan Khachatryan et al. SM Higgs Combination. *CMS Physics Analysis Summary, CMS-PAS-HIG-11-011*, 2011.
- [129] Serguei Chatrchyan et al. Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV. *Phys. Lett.*, B710 :26–48, 2012.

- [130] Serguei Chatrchyan et al. Observation of a new boson with a mass near 125 GeV. *CMS Physics Analysis Summary, CMS-PAS-HIG-12-020*, 2012.
- [131] Serguei Chatrchyan et al. Evidence for a new state in the search for the standard model Higgs boson in the H to ZZ to 4 leptons channel in pp collisions at $\sqrt{s} = 7$ and 8 TeV. *CMS Physics Analysis Summary, CMS-PAS-HIG-12-016*, 2012.
- [132] Georges Aad et al. Combined Measurement of the Higgs Boson Mass in *pp* Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett.*, 114 :191803, 2015.

Miscellaneous

- [133] ATLAS and CMS Collaborations. Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV. *ATLAS CONF-2015-044, CMS-PAS-HIG-15-002*, 2015.

Remerciements

Alors que l'exercice complexe des remerciements semble assidûment assumé dans les thèses de doctorat, les mémoires d'HDR en sont souvent dépourvues. Serait-ce à dire que la pudeur grandirait avec la maturité et freinerait l'expression de ces sentiments ? Néanmoins, il serait dommage de manquer l'occasion de remercier, qui somme toute se présente rarement, mais ne laisse jamais indifférent. Une des difficultés première de l'exercice est de tenter de n'oublier personne, ce qui est quasi-impossible, par conséquent je tiens par avance à présenter mes excuses à toute personne que j'oublie de nommer ici. La deuxième difficulté consiste à exprimer en quelques mots ce qui pourrait prendre bien plus d'espace, tout en étant suffisamment personnel pour exprimer ce que l'on pense vraiment, sans trop l'être dans un tel document public...

Par où commencer ? Tout d'abord je souhaite remercier les membres de mon jury qui ont tous accepté ce rôle avec enthousiasme : Olivier Drapier, Arnaud Lucotte, Emmanuel Perez, Ivica Puljak et Laurent Serin. Un grand merci pour vos rapports et commentaires, dont la teneur m'a fort touchée. J'ai eu l'occasion, à des degrés et sur des sujets bien différents, de collaborer avec chacun de vous et ce fut un réel plaisir à chaque fois ! En vrac et entre autres, je me rappelle avec tendresse les J(R)JCs, les masterclasses, Desy, Fermilab, le CERN, Split, des balles très légères tapées à l'aide de raquettes sur une table, les acrobaties sur un certain calorimètre à Argon liquide, des menus de *trigger* appétissants, des bruits de fond à estimer, les PTDR (*Physics Technical Design Reports* bien sûr, pas l'autre acronyme) soumis dans un état second à minuit moins une minute, les nuits blanches, les “*worries of the day*”, les conversations passionnantes souvent initiées par une question faussement insignifiante.

Merci à Réza Ansari et Marie-Christine Mignier pour leur disponibilité à l'égard du volet administratif de l'HDR à l'Université Paris-Sud.

Je suis immensément reconnaissante à mes relecteurs, Pierre Barrillon et Christophe Ochando, qui ont donné de leur temps pour améliorer de façon conséquente la qualité de ce manuscrit, et de la présentation de soutenance, par leurs commentaires percutants et pointus.

Merci au groupe CMS du LLR pour ces (déjà nombreuses) années partagées, pour la confiance accordée, et pour avoir accepté mon choix de prendre ce temps de rédaction, tout comme celui de prendre du temps de maman. Un ensemble original, harmonieusement com-

posé de personnalités toutes singulières, brillantes, et attachantes, dirigé avec dextérité par (en ordre chronologique), Ludwik Dobrzynski, à l’attention toujours chaleureuse, Philippe Busson, dont la discrétion et l’efficacité vont de pair avec sa grande humanité, et Yves Sirois, avec son enthousiasme et son impétuosité inaliénables, sa connaissance profonde de la physique et son intelligence habile. Merci.

Je me souviens de ma première venue au LLR, un jour de barbecue d’été, et de l’accueil si chaleureux par le légendaire sourire de Caroline Collard, qui allait sceller une grande amitié. De la “mise dans le bain” directe par les études sur le ECAL en campagne de tests en faisceaux, avec Pascal Paganini et Nicolas Regnault, qui m’ont accueillie dans leur bureau où régnait une ambiance très stimulante. Merci Pascal pour le partage enrichissant de bureau (à deux reprises) et nos “tours” de mauvaise humeur, nos heures de travail, et nos conversations passionnantes pendant toutes ces années, sur la physique et au-delà. Merci à Nicolas pour son expertise impressionnante sur le code, qui m’a bien aidée. Cette époque lointaine (quand la cantine de l’X était goûteuse) évoque également le souvenir de Gérard Milleret, sa rigueur diablement efficace, sa sempiternelle orange et ses lois si pertinentes.

Un peu plus tard, est venu le temps de l’émulation des *Trigger Primitives*, et notre fructueuse et faste collaboration avec Ursula Berthon, à qui j’adresse mes plus tendres pensées. Merci pour ces moments de réel partage et de confiance mutuelle, et l’amitié qui en a découlé (je ris toujours au veto que nous avons imposé à Pascal) ! Une pensée également pour Muriel Cerutti. Je me souviens notamment de collocations fort sympathiques avec Ursula et Muriel à l’occasion de déplacements. Une autre pensée pour l’expertise de Thierry Romanteau. Merci également Philippe Busson pour le temps que tu as patiemment passé à nous expliquer en détails les secrets de la chaîne de déclenchement du ECAL. Un grand merci Alexandre Zabi, toi qui as endossé à la suite de Pascal la responsabilité lourde du déclenchement ECAL au CERN, pour notre courte mais efficace collaboration, pour ton enthousiasme et ta résistance à toute épreuve, nos parties de rigolades nombreuses, ton aide, et ton amitié solide et précieuse (et un souvenir mémorable de notre périple à Key West). En continuant sur le thème du déclenchement, je remercie Thibaut Hennequin et Pauline Soulet qui ont fourni un travail remarquable sur l’étude des *triggers* pour l’analyse du boson de Higgs en quatre leptons. Pauline, une fraction non négligeable de tes résultats est exposée dans ce manuscrit, cela a été un plaisir et une satisfaction réels de collaborer avec toi. Merci à Nadir Daci, dont l’expertise et le document de thèse constituent une référence sur le sujet. Enfin, un grand merci Emmanuel Perez de ton soutien pour la défense des *triggers* essentiels à l’étude des performances des électrons.

Merci à la fine et solide équipe des électrons et du boson de Higgs en quatre leptons, qui a connu des collaborations variées ainsi que des phases bien différentes. Quel chemin parcouru depuis les pionniers et irréductibles, Claude Charlot, Ivica Puljak, Yves Sirois, Roberto Salerno, rejoints au fil des années par Christophe Ochando, Florian Beaudette, Jean-Baptiste Sauvan, et épaulés par toute une génération d’étudiants brillants, Federico Ferri, Chiara Rovelli, Jona Hampe, Arabella Martelli, Clémentine Broutin, David Sabès, Roko Pleština,

Marko Kovač, Mykhailo Dalchenko, Simon Regnard, Luca Cadamuro, Philipp Pigard, Toni Šćulac, par les non moins brillants ingénieurs en informatique, David Chamont, Ursula Berthon, Andrea Sartirana, Emilia Becheva, Arnaud Chiron ; sans oublier Paolo Meridiani, Nikola Godinovic, Nicola de Filippis, Damir Lelas, Chiara Mariotti, Nicola Amapane, Cristina Botta, Emmanuel Perez, Giacomo Ortona. En particulier, merci à Yves, pour son talent visionnaire et son énergie. À Claude, pour les innombrables heures passées ensemble à poser des questions, et y répondre, les nuits blanches, pour m’avoir fourni les clefs des électrons, pour nos débats animés, le co-encadrement de la thèse de Mykhailo, sans oublier la place dans l’amphi le 4 juillet 2012 ! Ivica, merci pour tous ces moments de connexion intellectuelle, ces autres de *rush* complètement fous, ta passion joyeuse, positive et communicatrice, ce plaisir à travailler ensemble, ton sens inné du compromis. Merci à Roberto, maestro du *slide* parfait, pour nos réflexions intenses et notre interaction efficace, je pense en particulier à ce mois d’août, le labo désert, notre trio de choc avec Ivica, et la création de l’ensemble de l’architecture et de l’algorithmie de l’analyse. Merci à Christophe Ochando, ninja à l’humour saisissant, pour notre collaboration si facile, nos partages, nos instants sushis, ton aide, ton écoute et ton amitié. Merci à Florian Beaudette, pour notre travail sur les électrons, pour ton aide et soutien sur le papier, pour ton don de faire sentir tes interlocuteurs mis en valeur. Je souhaite également remercier Chiara Rovelli pour notre superbe collaboration sur le papier, et Arabella Martelli pour nos échanges sur la quantité de matière du trajectographe. Enfin, merci à Jean-Baptiste Sauvan pour ton efficacité redoutable, ta réactivité à mes multiples demandes pour les figures du papier, et tes explications patientes et éclairantes sur la régression.

Une part non négligeable de mon temps a été consacrée pendant ces années à la “Communication” ou à la formation des jeunes (ou futurs) chercheurs : merci à Lison Bernet, Perrine Royole-Degieux, Caroline Collard, Auguste Besson, Raphaël Granier de Cassagnac, Nicolas Regnault, Pierre Barrillon, Ludwik Dobrzynski, Philippe Miné, Olivier Drapier, Christophe Ochando, Christophe Thiebaut, François Arleo, Sylvaine Pieyre pour tous ces échanges enrichissants.

Il est assez étrange d’être le témoin du *turn-over* dans un bureau. Le mien a vu des collaborateurs nombreux, et sympathiques : merci à Pascal Paganini, Nicolas Regnault, Stéphane Bimbot, Jona Hampe, Evelyne Delmeyer, Andrea Benaglia, Lorenzo Bianchini, Colin Bernet, Pauline Soulet, Roko Pleština, Michal Bluj, Nadir Daci, Ivo Naranjo Fong, Luca Mastrolorenzo (merci pour les pauses souvent bien drôles !), Philipp Pigard, Yurii Antropov, pour avoir construit l’ambiance de ce bureau.

Merci également à ceux avec qui je n’ai pas directement collaboré, mais dont je n’ai pas moins apprécié les interactions : Lamia Benhabib, Émilien Chapon, Alice Florent, Thomas Strebler, Alain Bonnemaïson, Berrie Giebels, Pietro Govoni, et Olivier Davignon. Merci Olivier pour nos pauses cigarettes toujours agréables, nos discussions souvent animées, les parties de ping-pong endiablées, ta bienveillance et ton aide généreuse.

Merci au LLR, à tous ses membres qui en font la singularité, à Jean-Claude Brient qui le

dirige (merci pour le petit cadeau estampillé LLR) et aux services qui restent dans l'ombre mais sans qui rien ne pourrait arriver. Merci au service administratif dirigé par Thu Bizat, et en particulier à Brigitte Maroquesne, résistante à toutes les demandes mêmes les plus farfelues, également les dépannages de cigarettes, Marie-Thérèse Théodora qui sait répondre à toutes les interrogations administratives, Nora Aissous pour son aide dans les méandres de l'X, Sylvaine Pieyre et Élodie Dubois pour leur efficacité souriante. Merci à l'équipe de l'informatique, et en particulier à Emilia Becheva, Éric Beyer, Pascale Hennion, Michel Lastes, Andrea Sartirana, Igor Semenouk, Julien Tugler, Gilles Grasseau, pour leur soutien lors de demandes informatiques en tout genre. Merci Yannick Geerebaert pour les convois dans ta voiture lors des séances défolement au ping-pong. Merci également à Hamid Khaled pour son aide et son esprit pratique ! Enfin, je suis reconnaissante à tous ceux présents le jour de la soutenance ; votre aide, votre attention et vos mots de congratulations m'ont touchée.

Si tant est que l'on peut remercier les choses ou concepts, je souhaite remercier la page CRISCO des synonymes (<http://www.crisco.unicaen.fr/des/synonymes/>), demeurée ouverte tout au long de la rédaction de ce document. Également, la danse et les danses à deux, et en particulier la danse orientale et le tango argentin pour l'équilibre et le moyen d'expression qu'ils m'apportent. Merci à Selma Conforti pour son amitié, ainsi qu'aux copines du cours, pour ces moments de rigolade à l'observation d'un œil critique de nos prestations dans le miroir. Merci à David Chamont pour le partage autour du tango. Je suis également reconnaissante à Sylvie Pradelle pour son aide suivie et son guidage en douceur vers le lâcher prise.

Sur un registre plus personnel, j'adresse tout d'abord une pensée émue à ceux qui ont quitté ce monde. Je suis reconnaissante à mes amis de toujours (ou presque) : Célia d'Abri-geon, Valérie Gautrot, Nicolas Ponthieu pour leur amour à toute épreuve. Un grand merci aux familles Barrillon & Chaud qui m'ont si naturellement fait une place parmi eux, et en particulier Gaëlle Cloarec, Geneviève et Michel Barrillon pour leur affection chaleureuse. Merci à toute ma famille, et en particulier à ma soeur pour sa perspicacité, sa compréhension et sa bienveillance, et à mes parents qui ont su m'enseigner l'amour et le don de soi, sans compter. Mes derniers remerciements s'adressent à Pierre Barrillon, qui sait m'entourer de sa tendresse, de sa gentillesse, de sa confiance, de sa compréhension, tout en me faisant rire, à notre équipe parentale de choc, et à nos enfants, Nathanaël et Siméo qui savent pimenter et faire pétiller nos vies...

*I've learned that people will forget what you said,
people will forget what you did,
but people will never forget
how you made them feel.*
Maya Angelou

